Technical Report 1335

Framework for Instructional Technology: Methods of Implementing Adaptive Training and Education

Paula J. Durlach
Randall D. Spain
U.S. Army Research Institute

January 2014

United States Army Research Institute
for the Behavioral and Social Sciences

U.S. Army Research Institute
for the Behavioral and Social Sciences

Department of the Army
Deputy Chief of Staff, G1

Authorized and approved for distribution:

MICHELLE SAMS, Ph.D.
Director

Technical review by

Peter S. Schaefer, U.S. Army Research Institute
Cheryl I. Johnson, U.S. Army Research Institute

NOTICES

DISTRIBUTION: Primary distribution of this Technical Report has been made by ARI. Address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: DAPE-ARI-ZXM, 6000 6th Street, (Bldg 1464 / Mail Stop 5610), Ft. Belvoir, Virginia 22060-5610.

FINAL DISPOSITION: Destroy this Technical Report when it is no longer needed. Do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this Technical Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

<table>
<tr><td colspan="5">REPORT DOCUMENTATION PAGE</td><td colspan="2">Form Approved
OMB No. 0704-0188</td></tr>
<tr><td colspan="2">1. REPORT DATE (DD-MM-YYYY)
January 2014</td><td colspan="3">2. REPORT TYPE
Final</td><td colspan="2">3. DATES COVERED (From - To)
June 2011 – April 2013</td></tr>
<tr><td colspan="5" rowspan="3">4. TITLE AND SUBTITLE
Framework for Instructional Technology:
Methods of Implementing Adaptive Training and Education</td><td colspan="2">5a. CONTRACT NUMBER</td></tr>
<tr><td colspan="2">5b. GRANT NUMBER</td></tr>
<tr><td colspan="2">5c. PROGRAM ELEMENT NUMBER
633007</td></tr>
<tr><td colspan="5" rowspan="3">6. AUTHOR(S)
Paula J. Durlach
Randall D. Spain</td><td colspan="2">5d. PROJECT NUMBER
A792</td></tr>
<tr><td colspan="2">5e. TASK NUMBER
363</td></tr>
<tr><td colspan="2">5f. WORK UNIT NUMBER</td></tr>
<tr><td colspan="5">7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute
for the Behavioral and Social Sciences
6000 6th Street (Building 1464 / Mail Stop 5610)
Fort Belvoir, VA 22060-5610</td><td colspan="2">8. PERFORMING ORGANIZATION REPORT NUMBER
Technical Report 1335</td></tr>
<tr><td colspan="5" rowspan="2">9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Institute
for the Behavioral & Social Sciences
6000 6th Street (Building 1464 / Mail Stop 5610)
Fort Belvoir, VA 22060-5610</td><td colspan="2">10. SPONSOR/MONITOR’S ACRONYM(S)
ARI</td></tr>
<tr><td colspan="2">11. SPONSOR/MONITOR’S REPORT NUMBER(S)
Technical Report 1335</td></tr>
<tr><td colspan="7">12. DISTRIBUTION/AVAILABILITY STATEMENT: Distribution Statement A: Approved for public release; distribution unlimited.</td></tr>
<tr><td colspan="7">13. SUPPLEMENTARY NOTES
Contracting Officer’s Representative and Subject Matter Experts: Paula J. Durlach & Randall D. Spain</td></tr>
<tr><td colspan="7">14. ABSTRACT
In adaptive instructional environments, instructional interventions and/or content are adapted to an individual learner’s competence level, goals, or other characteristics. The intention behind adaptation is to maintain the optimal level of challenge for each individual student, to provide support, and to correct misconceptions. This report provides a framework in which to consider various technology-based adaptive instructional techniques. The Framework for Instructional Technology (FIT) lays out four categories of adaptive techniques: Corrective Feedback, Support, Micro-sequencing, and Macro-sequencing. Under each category, FIT specifies five “levels” or approaches to adapting that correspond to degree of adaptive sophistication and complexity of implementation. With few exceptions, evidence supporting the use of higher levels of adaptation is lacking. This is because the systematic comparison of different implementation approaches has not been conducted. The report provides recommendations for combining different levels of FIT with different levels of interactive multimedia instruction. FIT can be used to clearly describe adaptive behaviors and to guide future research investigating the efficacy of different adaptive instructional interventions.</td></tr>
<tr><td colspan="7">15. SUBJECT TERMS
Adaptation, Instructional technology, Adaptive training, Feedback, Sequencing, Intelligent tutor, Learning</td></tr>
<tr><td colspan="3">16. SECURITY CLASSIFICATION OF:</td><td rowspan="2">17. LIMITATION OF ABSTRACT
Unlimited
Unclassified</td><td rowspan="2">18. NUMBER OF PAGES
56</td><td colspan="2" rowspan="2">19a. NAME OF RESPONSIBLE PERSON
Dorothy Young
703-545-2316</td></tr>
<tr><td>a. REPORT
Unclassified</td><td>b. ABSTRACT
Unclassified</td><td>c. THIS PAGE
Unclassified</td></tr>
</table>

Technical Report 1335

Framework for Instructional Technology: Methods of Implementing Adaptive Training and Education

Paula J. Durlach and Randall D. Spain
U. S. Army Research Institute

Orlando Research Unit
Joan H. Johnston, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
6000 6th Street, Bldg. 1464
Fort Belvoir, VA 22060

January 2014

FRAMEWORK FOR INSTRUCTIONAL TECHNOLOGY: METHODS OF IMPLEMENTING ADAPTIVE TRAINING AND EDUCATION

EXECUTIVE SUMMARY

Research Requirement:

As outlined in the Army Learning Concept 2015 (ALC 2015), Army training and education is undergoing a transformation to a learner-centric model. As this occurs, learning outside the classroom will play an increasingly key role. Innovative learning technologies and methods will be required to make self-directed learning effective and efficient. One of the items in the ALC 2015 Action Plan is: identify state-of-the-art adaptive training and digital tutor capabilities, and develop standards, protocols, and guidance on employing these capabilities in interactive multimedia instruction (IMI) modules.

Procedure:

Adaptive instructional technology can be implemented in a myriad of ways. There does not appear to be a standard terminology for describing the various possible implementations. Digital tutors are considered the epitome of adaptive instructional technology; however, different tutors are implemented in different ways. Moreover, there are adaptive tactics that may be just as effective as those typically used by digital tutors, but less technically difficult to implement. In this report, we examined the literature on adaptive instructional technology and created a framework that allows comparison of many of the potential adaptive techniques. We also considered how the various techniques could be used in combination with different levels of IMI.

Findings:

The Framework for Instructional Technology (FIT) lays out four categories of adaptive techniques: Corrective Feedback, Support, Micro-sequencing, and Macro-sequencing. Corrective Feedback covers methods for responding to incorrect responses. Support covers methods of providing support, cues, hints, and prompts. Micro-sequencing covers methods of remediation, given that a performance criterion has not been met. Finally, macro-sequencing covers how the sequencing of lessons or modules is determined (once mastery has been met in one module, what next?). Under each category, FIT specifies five “levels” or approaches to adapting. These levels roughly correspond to the degree of adaptive sophistication and complexity of implementation. Whereas for corrective feedback, support, and micro-sequencing, the instructional designer must select one level to use, macro-sequencing is somewhat different. For macro-sequencing, the levels are not mutually exclusive and can be combined in various ways. For each category, we examined the status of the empirical evidence as to whether it would justify selection of a more complex over a less complex form of implementation. In the case of Corrective Feedback, the evidence quite clearly justifies recommending the use of error-sensitive feedback; however, for the other categories, evidence supporting the use of higher levels of adaptation is lacking. The appropriate systematic comparison of different

implementation approaches has simply not been conducted. The report provides recommendations for combining different levels of FIT with different levels of IMI (see Tables 8 and 9).

Utilization and Dissemination of Findings:

There are two primary uses of FIT. The first is to serve as a guideline to instructional designers or for those procuring instructional technology. FIT can be used as a method of clearly specifying the adaptive behaviors that a new instructional system should manifest. The second use is to serve as a framework for future research on the efficacy of different adaptive techniques. As previously mentioned, the evidence to justify the use of certain adaptive methods (such as Level III Adaptive Content Micro-sequencing—typically used by digital tutors), over other less complex approaches (e.g., Level I or II Micro-sequencing) is lacking. This is because much empirical research in educational technology compares the effects of multiple adaptive features against no adaptive features. This makes it impossible to determine the relative contribution of the different features, when beneficial effects on learning outcomes are achieved. Research should be conducted holding constant all factors except for a single level of a single FIT category, in order to determine the relative efficacy of the two levels.

An earlier, slightly different version of FIT was presented at Applied Digital Human Modeling 2012, San Francisco, as part of the International Conference on Applied Human Factors and Ergonomics in July 2012, and published as Durlach and Spain (2012). FIT was also presented in a talk at The Advanced Distributed Learning Initiative's iFest, August 2012.

FRAMEWORK FOR INSTRUCTIONAL TECHNOLOGY: METHODS OF IMPLEMENTING ADAPTIVE TRAINING AND EDUCATION

CONTENTS

Page

INTRODUCTION 1

EXISTING TERMINOLOGY DESCRIBING ADAPTIVE TECHNOLOGY-BASED INSTRUCTION 2
- Adaptive Instruction 2
- Macro and Micro Adaptation 3
- Inner-Loop and Outer-Loop 3
- Granularity and Steps 4

THE FRAMEWORK FOR INSTRUCTIONAL TECHNOLOGY 5
- Corrective Feedback 7
- Support 11
- Micro-sequencing 15
- Macro-sequencing 23
- Sequencing vs. Branching 27
  - Student control of sequencing within a lesson or module 27
  - Branching scenarios 27

THE RELATION BETWEEN FIT AND IMI 29
- IMI Level One – Passive 29
- IMI Level Two – Limited Participation 29
- IMI Level Three – Complex Participation 31
- IMI Level Four – Real-time Participation 33

SUMMARY 34

REFERENCES 37

ACRONYMS 44

TABLES

Page

TABLE 1: TYPES OF LOCALLY-ADAPTIVE CORRECTIVE FEEDBACK 9

TABLE 2: THE FIVE LEVELS OF CORRECTIVE FEEDBACK SPECIFIED IN FIT......... 11

TABLE 3: THE FIVE LEVELS OF SUPPORT SPECIFIED IN FIT 14

TABLE 4: AN ILLUSTRATION OF HIERARCHICALLY ARRANGED LEARNING OBJECTIVES. 17

TABLE 5: AN ILLUSTRATION OF THE ASSESSMENT RESULTS OF TWO HYPOTHETICAL STUDENTS. 18

TABLE 6: THE FIVE LEVELS OF MICRO-SEQUENCING SPECIFIED IN FIT 20

TABLE 7: THE FIVE LEVELS OF MACRO-SEQUENCING SPECIFIED IN FIT 25

TABLE 8: TYPICAL FIT LEVELS ASSOCIATED WITH ITS, AND RECOMMENDED FOR DIFFERENT LEVELS OF IMI 34

TABLE 9: FIT SUMMARY TABLE 35

FIGURES

Page

FIGURE 1. SCHEMATIC ILLUSTRATION OF THE INNER LOOP (DASHED LINES) AND THE OUTER LOOP (SOLID LINES), TYPICAL OF ITS............................ 4

FIGURE 2. ILLUSTRATION OF THE TWO DECISIONS A SYSTEM MUST MAKE WHEN MASTERY LEARNING IS IMPLEMENTED.. 6

FIGURE 3. ILLUSTRATION OF WHERE FEEDBACK, SUPPORT, MICRO-SEQUENCING, AND MACRO-SEQUENCING OCCUR DURING THE INSTRUCTIONAL PROCESS. 7

FIGURE 4. EXAMPLE OF A MULTIPLE-CHOICE QUESTION. THE CORRECT ANSWER IS A POISONOUS FALSE MOREL. 10

FIGURE 5. LEVELS 0, I, AND II OF MICRO-SEQUENCING IN PANELS A, B, AND C, RESPECTIVELY.. 20

FIGURE 6. EXAMPLE OF HOW MICRO-SEQUENCING CAN PRODUCE TWO DIFFERENT PATHS IN SHOOT/NO SHOOT DECISION TRAINING.............. 22

FIGURE 7. SCHEMATIC ILLUSTRATION OF STUDENT CONTROL OVER SEQUENCING DURING A LESSON. 28

FIGURE 8. SCHEMATIC ILLUSTRATION OF A NOTIONAL BRANCHING SCENARIO WITH THREE OPTIONS AT EACH DECISION POINT................ 28

FRAMEWORK FOR INSTRUCTIONAL TECHNOLOGY: METHODS OF IMPLEMENTING ADAPTIVE TRAINING AND EDUCATION

INTRODUCTION

Over the past two decades, digital learning environments have extended traditional methods of instruction within the Department of Defense's (DOD) educational and training strategy. Digital learning environments provide not only education, but also training and skill application to learners distributed across the globe in controlled and safe environments, using simulations and games. More recently, the DOD has embraced the idea that interactive multimedia instruction (IMI) should be more learner-centric. One of the implications is that future technology-enabled learning environments should adapt themselves to the needs, understanding, and experience of the learner. In addition, those environments should integrate advances in learning science and evidence-based pedagogical methods (Fautua, Schatz, Reitz, & Killilea, 2012; TRADOC, 2011).

Adapting instructional content or strategy to the needs and ability of the learner can be done in many different ways. Adaptive interventions are possible whenever there is information about or input from the learner, collected either prior to or during the course of learning. Both the data used and the system's adaptive response involve design decisions such as: How will the system capture the necessary data about the student? What data are required and how should they be interpreted? How should the system respond to the different possible interpretations? When should it respond—during or after exercise completion? The range of possible adaptive behaviors is large; but most systems only incorporate a few. Those that are incorporated and how they are implemented is typically based on a combination of knowledge of learning science, craft, intuition, time, and resources. The main purpose of this report is to lay out a framework that outlines a range of possibilities. Our focus is primarily on the adaptive behavior of the instructional system –i.e., how the system responds to different student inputs and characteristics. This necessitates some discussion of the student information considered; however, a detailed discussion of the collection and processing of data used to inform adaptive decisions is beyond the scope of this report. The interested reader is directed to Behrens, Mislevy**,** Dicerbo, and Levy (2010), Desmarais and Baker (2012), and Shute and Kim (2013).

Adaptive design decisions are implemented for the potential benefits they may produce for learning speed, learner satisfaction, mastery and transfer; but they almost invariably increase the resources required for implementation. Consequently, there is a practical need to consider return on investment when selecting whether and how to adapt. To address this need, discussion of the framework includes comments regarding the existence of empirical evidence concerning learning outcomes. To foreshadow the discussion, the empirical evidence is sparse. Therefore an additional use of the framework should be to guide future empirical research by outlining the critical conditions that should be compared.

EXISTING TERMINOLOGY DESCRIBING ADAPTIVE TECHNOLOGY-BASED INSTRUCTION

Before presenting our framework, we will briefly define several key terms and review various existing frameworks that impose some organization on the range of potential adaptive behaviors a technology-enabled learning environment might display.

Adaptive Instruction

In this report we refer to adaptive instruction as training or education in which content, feedback, scaffolding, or support is tailored to an individual learner's aptitudes, learning preferences, or styles, either before instruction, or in real-time during instruction, with the aim of enhancing learning outcomes (Landsberg, Astwood, Van Buskirk, Townsend, Steinhauser, & Mercado, 2012; Shute & Zapata-Rivera, 2008). Human tutoring is the epitome of adaptive instruction. An ideal human tutor combines what they know about the student, effective instruction, and the subject to dynamically adapt instruction based on the perceived needs of the student. The tutor may give hints or guiding questions to stimulate dialog about the topic; they may give different kinds of feedback based on how close or far away a student's response is from being correct; likewise, they may use knowledge about the student's strengths and weaknesses to assign remedial content or exercises. The education benefits of one-on-one tutoring are clear: students working with a good tutor obtain higher achievement levels compared to students who receive instruction in a conventional classroom (Bloom, 1984). For this reason, many educational researchers and instructional technologist have looked for ways to replicate the benefits of one-on-one tutoring in adaptive instructional environments.

Intelligent tutoring systems (ITS) are the closest technological analogue of one-to-one human tutoring. These systems differ from traditional forms of computer-mediated instruction in that they reason about *what*, *when*, and *how* to teach while providing one-to-one individualized instruction (Woolf, 2009). These decisions are typically arrived at through use of various artificial intelligence models, such as a: student model, which maintains an estimate of the student's knowledge, skills, and abilities; domain model, which contains information about the training topic and how experts solve problems within the domain; and pedagogical model, which uses information from the student and domain model to prescribe different tutoring strategies. It is important to note not all ITS are created equal - some are more "intelligent" than others, meaning they encode more information and logic about the student, domain, and tutoring strategy in these models than others (Shute & Psotka, 1996). For example, some ITS may only provide feedback and hints based on an isolated student action or input without taking information from the student model into account. Others may maintain a rich database of student interactions and use this information to prescribe tailored instructional interventions.

Ideally, anything and everything that is known about the student, and that could influence their learning, should be considered when making adaptive instructional decisions (i.e., how, what, and when to make an adaptation). Of course, in reality, the full gamut of information is not available, and decisions are made on the basis of a simplified model of reality, based on the information that is available. They are also made on the basis of feasibility, such as whether software can be made intelligent enough to detect student state and respond appropriately, or

whether instructional environment actually has the time and resources to implement an adaptive strategy. For this reason, in this paper we discuss various ways and approaches in which adaptive instruction can be achieved; some approaches are relatively low on the continuum of adaptive training and range from using assessment scores to determine which training module a student should take next to tailoring the training content based on a learner's preferences. Others are more complicated and rely on robust student modeling to deliver prescribed feedback, hints, and support to students based on a contextual understanding of the student's needs. In the following sections, we review several existing rubrics that categorize the different behaviors of adaptive instructional environments. These concepts and terms are important for understanding the range of possible adaptive behaviors.

Macro and Micro Adaptation

Shute (1993) and Park and Lee (2004) describe adaptive systems as being either macro-adaptive or micro-adaptive, although a single system could contain aspects of both. Macro-adaptation uses pre-task measures or historical data to adapt content before the instructional experience begins. There are three general macro-adaptive approaches: Adaptation-as-Preference, Role Adaptation, and Mastery. With Adaptation-as-Preference, learner preferences are collected before training and this information is used to provide personalized training content. For example, it may determine whether a student watches a video or reads; or whether examples are given with surface features about sports, business, or the military. With Role Adaptation, trainees select their role (e.g., tank driver vs. tank gunner) and are then presented with different subsets of the content according to the role's learning objectives. With Mastery macro-adaptation, a pretest determines the starting point of instruction (allowing already mastered material to be skipped).

Micro-adaptive systems respond to trainee data in a dynamic fashion. These systems perform on-going adaptations during the learning experience, based upon the performance of the learner or other behavioral assessment (e.g., frustration, boredom). They may use a pattern of response errors, response latencies, and/or emotional state to identify student problems or misconceptions. Micro-adaptive moves may be aimed at correcting specific errors and their underlying cause. They may also provide support, such as giving hints or encouragement, or by eliciting reflection.

Inner-Loop and Outer-Loop

VanLehn's (2006) description of the behavior of ITS has popularized the use of the terms inner and outer-loops. The inner-loop is responsible for providing within problem guidance or feedback, based on the most recently collected input. The outer loop is responsible for deciding what task a student should do next, once a problem has been completed. Figure 1 illustrates the functioning of the inner (dashed lines) and outer (solid lines) loops. In this figure, the student has responded correctly to the item, and received confirmation with the "Good" feedback message via the inner loop. In ITS, data collected during inner-loop student-system interactions are used to update a student model representing the student's evolving competence. Comparisons of the student model with a domain model drive the outer loop decision process, tailoring selection of the next problem. Selection of the problem is based on the student's

relative grasp of the different learning objectives targeted in that module of instruction, with the aim of keeping the student at the right level of challenge. This necessitates problems being assigned different weights for the knowledge components (i.e., concepts, facts, principles, or procedures) required to solve them. It is this "intelligent" selection of the next problem that sets ITS apart from less adaptive systems, which may present problems in a fixed order, or in a random order. Reflecting back on the macro- and micro-adaptive characterization, the inner- and outer-loop processes are both micro-adaptive, because they use data collected during the course of learning.

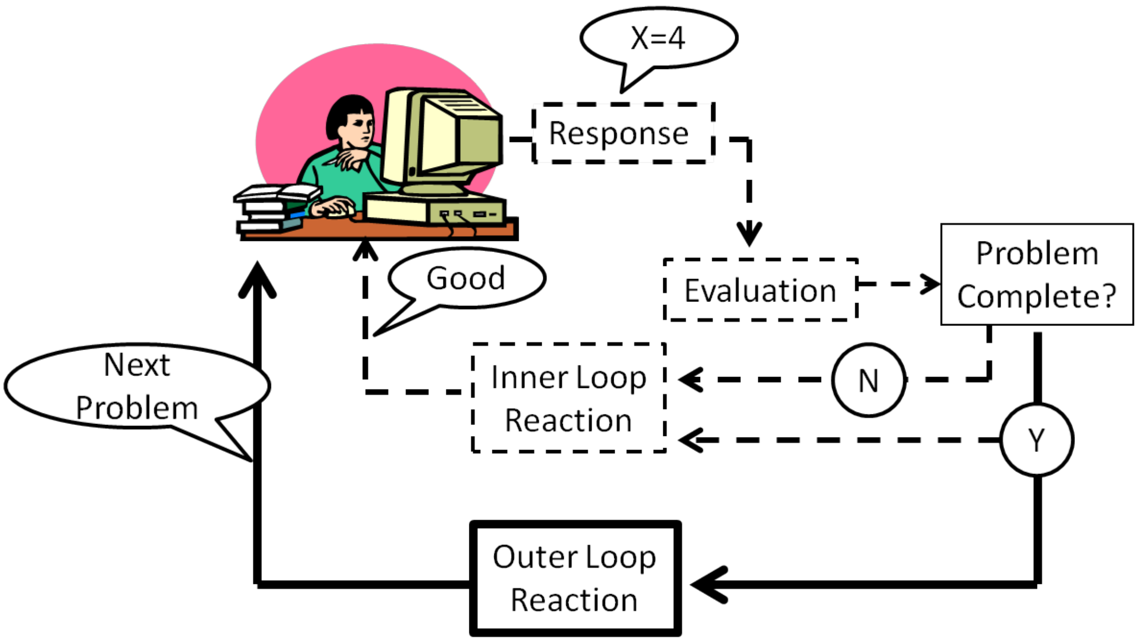

Figure 1. Schematic illustration of the inner loop (dashed lines) and the outer loop (solid lines), typical of ITS.

Granularity and Steps

ITS and other tutoring systems can also be characterized by the granularity of student interaction with the system. VanLehn (2011) distinguished four levels of granularity: none, answer-based, step-based, and substep-based. These levels correspond to the number of opportunities students have to interact with the tutoring system when solving a problem. With none, the student performs the required problem solving activity (multiple steps), with no system intervention. With answer-based, the student completes a problem and then submits the solution. Once the solution is submitted (or an activity is completed, say, for simulation-based training), the tutor provides feedback on whether the final answer is correct or not. If the answer is correct the student may be presented with a new problem. If the answer is incorrect the student may get another chance to resubmit his or her answer (or complete the task). Systems like these generally lack an inner loop or what Heffernan, Koedinger and Razzaq (2008) refer to as the knowledge search loop. These systems may assign the student a task, collect the student's answer, give them feedback on their answer, and continue either by giving the student another chance to solve the problem or by going on to a new one; however, they do not provide step-based instructional

support (i.e., they do not assist the student in identifying *why* an answer was incorrect). Systems that lack an inner loop are sometimes referred to as learning content management systems, computer-aided instruction, or web-based homework systems (VanLehn, 2011).

Step-based and sub-step-based systems represent a finer level of granularity. With step-based systems, the student can receive guidance or feedback on each step of the solution. This is enabled by repetitive iterations of the inner loop, one for each step, until the problem is complete (VanLehn, 2011). The interface is typically designed specifically to support the student entering steps, and students may be able to request hints to get advice at each specific step. After each step they may receive some feedback. The granularity is finer because student's entries and the system's hints and feedback refer to a relatively smaller application of knowledge (the step), compared with answer-based systems (VanLehn, 2011). A step-based system can also be designed to provide feedback only after a complete answer is submitted. It is still considered step-based if the steps of the problem are reviewed in the feedback, enabling the student to see where their errors were committed (if any). Thus, a scenario-based instructional simulation with an after action review that presents what the student did right and what they did wrong, would be analogous to a step-based system, whereas one that simply gave them a score would be analogous to an answer-based system.

The level of support given in a substep-based system is even finer than step-based. Substep-based systems tend to use natural language interfaces, enabling a student-system dialog on each step; for example, a pedagogical agent may follow up a partially correct step asking the student a probing question, initiating a multi-turn dialog that results in the student analyzing their input and correcting it. In terms of effectiveness, VanLehn, 2011) suggested that substep-based and step-based systems were about equally effective, but more effective than answer-based, which in turn were more effective than none.

THE FRAMEWORK FOR INSTRUCTIONAL TECHNOLOGY

The preceding discussion illustrated that there are different varieties of micro-adaptation. These include methods for supporting student learning during the problem solving process, as well as methods for selecting what the next problem should be. The Framework for Instructional Technology (FIT)[1] aims to better differentiate potential implementations of these processes, and also to add the notion of between-module adaptation. FIT is organized around three important pedagogical tactics: corrective feedback, support, and mastery. Each of these is explained in greater detail later; for a primer, corrective feedback refers to post-response information that aims to help students repair misconceptions in their thinking or behavior. Support refers to any type of instructional scaffolding (e.g., hints, pumps, instructional diagrams, etc.) that is used to help a student master a concept or topic. Mastery refers to a pedagogical technique that tailors the instructional content to the student's current level of understanding; it requires students to demonstrate mastery of an exercise or topic before proceeding to the next, more advanced one. As a technique, mastery learning involves two basic decisions, as illustrated in Figure 2: What to do when a student reaches mastery, and what to do when a student fails to reach mastery. FIT attempts to parse these decision-points into two unique tactics: micro-sequencing and macro-

[1] As presented here, FIT is a slightly modified and improved version of the framework as originally presented in Durlach and Spain (2012).

sequencing. Micro-sequencing determines what to do next when a student needs remediation; macro-sequencing determines what to do once current topic mastery has been attained.

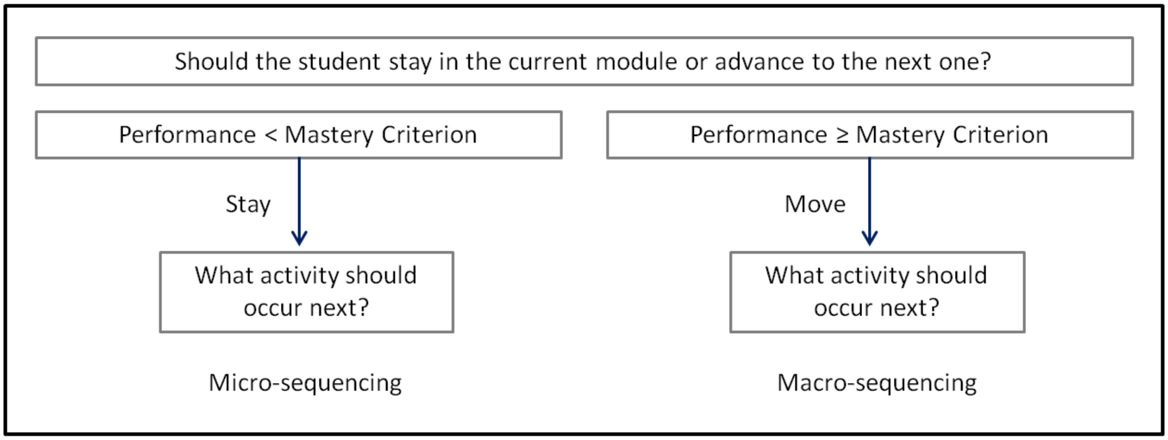

Figure 2. Illustration of the two decisions a system must make when mastery learning is implemented. When performance is below mastery (left) the system has to decide what learning activity should occur to bring the student closer to mastery. When performance is at or above mastery (right) the system has to decide what module to present next.

Figure 3 illustrates where these four tactics (feedback, support, micro-sequencing, and macro-sequencing) occur during the course of learning, in a schema analogous to Figure 1. Inspection of Figure 3 shows that feedback and support are inner-loop processes, whereas task selection, which was analogous to VanLehn's (2006) outer loop, is now controlled by micro-sequencing. More specifically, micro-sequencing controls the selection of activities within a lesson or module when mastery has not been reached. To this structure FIT adds a third loop called macro-sequencing. Macro-sequencing is like micro-sequencing, but at a higher level of analysis. Given that a student has mastered the learning objectives covered by a particular module of instruction, the macro-sequencing decision determines what module he or she should do next. In summary, micro- and macro-sequencing have to do with the sequencing of events, either within a lesson (micro) or across lessons (macro). Corrective feedback and support are more granular and provide within problem guidance, either at the answer, step, or sub-step level.

For each of these four tactics (corrective feedback, support, micro-sequencing, and macro-sequencing) FIT presents a range of potential ways that each can be implemented in technology-based instructional systems. Specifically, for each of these tactics, FIT lays out five different "levels" of adaptivity. These levels roughly correspond to the complexity of the underlying data structures needed to support the implementation, although in some cases, different methods of implementation do not fall along such a clear continuum (macro-sequencing in particular).

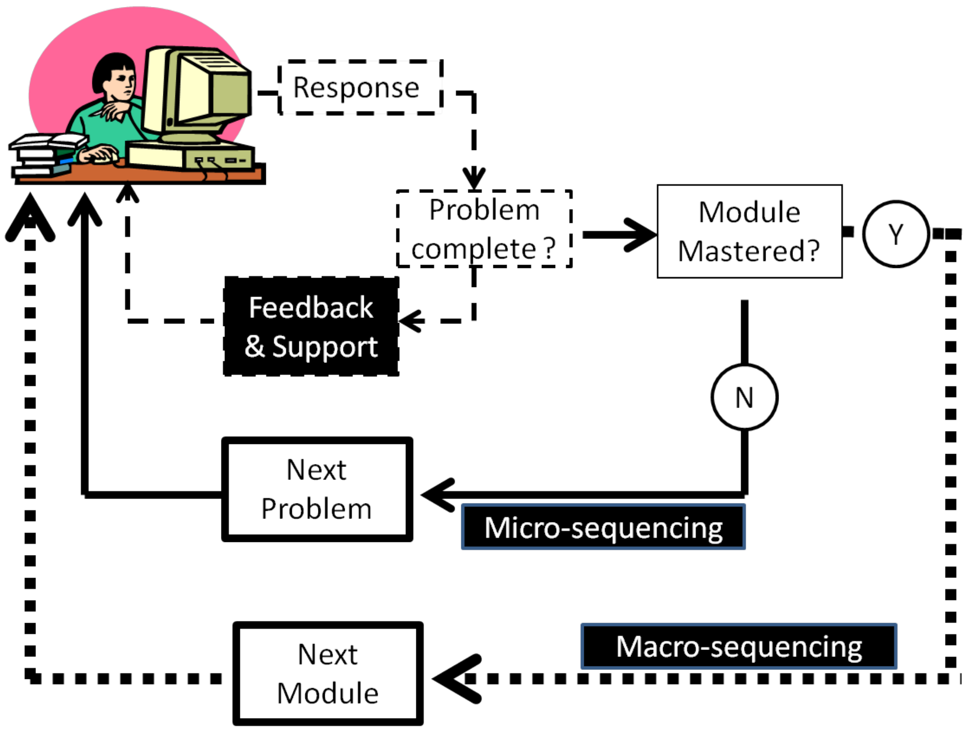

Figure 3. Illustration of where feedback, support, micro-sequencing, and macro-sequencing occur during the instructional process.

Corrective Feedback

It is well accepted that feedback is important for learning (Azevedo & Bernard, 1995; Hays, 2005; Jaehnig & Miller, 2007), yet there is still some debate concerning the most effective methods for providing it (e.g., see Hattie & Timperley, 2007; Hays, Kornell, & Bjork, 2013; Kluger & DeNisi, 1998; Schmidt & Bjork, 1992). It may be that the effectiveness of feedback depends on the type of task or the current state of student knowledge or ability (Van Dijk & Kluger, 2011). For example, Forbes-Riley and Litman (2011) found that adapting feedback content on the basis of student certainty (as well as accuracy) resulted in better post-test performance (compared with when feedback content was based on accuracy alone). Specifically, students who received different feedback content when they were correct, but uncertain, compared with when they were correct and certain, performed better on a posttest than students who got the same (positive) feedback for every correct response.

Feedback providing knowledge of results (both positive and negative) can be beneficial. Positive feedback can be motivating and reduce uncertainty. A critical role of feedback is to support the student in comparing his or her own performance with what good performance looks like, and enabling students to use this information to close that gap (Sadler, 1989). In general, the more fine-grained the feedback, the more likely it is to fit these requirements. A summary grade provides an evaluation, but does not go very far in helping the student "close the gap." (VanLehn, 2011) suggested that providing feedback in a way that allows students to self-correct

may be the most effective method of enhancing learning outcomes. Because we are concerned with "closing the gap," FIT concentrates on corrective feedback.

FIT is concerned with explicit corrective feedback, as opposed to the implicit feedback. Explicit feedback refers to information a student receives from an external source such as an instructor, system, or artificial agent that intends to inform the student how he or she performed on a task (Narciss, 2008). Receiving a message from an instructional environment about why an answer is incorrect is a form of explicit feedback. Implicit feedback, on the other hand, refers to information inherent in a scenario, task, or environment that may be used to determine the accuracy of a student's decision or input (Narciss, 2008). There is no agent or system that explicitly tells the student if his or her actions were correct or incorrect; rather this must be inferred based on changes to the environment or the goals of the task. We believe it is important to distinguish between these two forms of feedback to better understand the differences between interactive and adaptive instructional environments.

Both interactive and adaptive instructional environments alter their behaviors on the basis of student input. As we see it, the distinction between them concerns whether the system reaction is based on an evaluation of the student input (adaptive) or not (interactive). For example, in a driving simulator, stepping on the brake slows down the apparent motion of the vehicle in the virtual world, without making any evaluation of whether stepping on the brake was the correct thing to do. This is an example of interaction, but not adaptation. This type of interaction provides implicit feedback, which can be instructionally beneficial if it is noticed and interpreted by the student as resulting from his or her input; but the implicit feedback is generated by a vehicle simulation model that is indifferent to the correctness of stepping on the brake. An instructionally adaptive reaction, in contrast, would be one that evaluates the student input and produces a response based on that evaluation. In the current example, an adaptive response would be a feedback message praising the student for slowing down under low visibility conditions. According to FIT, any explicit feedback is a form of adaptation. As we will explain in the next section, however, it can be given in a variety of ways, depending on the richness of the information used for evaluation and/or provided in the feedback.

Table 1 presents steps on a continuum of corrective feedback in which error diagnosis and corrective information increase as one reads down the table. It illustrates how different forms of corrective feedback can be applied if a student were to err on the question posed in Figure 4. As shown in Table 1, the simplest explicit feedback is "minimal feedback." Minimal feedback simply indicates whether a response is correct or incorrect. In Sadler's terms (1989), it tells students whether there is a gap, but it does not tell them how to close it. When accuracy data are supplemented with more explanatory information, the feedback becomes more effective (e.g., Billings, 2012). More rich than minimal feedback, correct answer feedback tells the student what the correct response should be. This may be supplemented further with an explanation as to why it is correct (correct answer explanation). If the explanation also takes into account how the student erred, it is referred to as error-sensitive feedback. Error-sensitive feedback tells the student why his or her response is wrong, with the intention of repairing the error, and is therefore the most aligned with Sadler's prescription from among the options listed. With error-sensitive feedback, if a student incorrectly chooses B instead of A on a question, the feedback would explain why B is incorrect and the difference between A and B. For a different

student erring by selecting C, the feedback would explain why C is incorrect and the difference between A and C.

Table 1. ***Types of Locally-Adaptive Corrective Feedback***

Effectiveness	**Feedback type**	**Description**	**Example Feedback if student responds b. True Morel**	**Example Feedback if student responds c. Coral Fungus**
Low	Minimal Feedback	Tells whether a response was correct or incorrect	Incorrect	Incorrect
Low	Correct Answer	Provides correct answer	No, this is a False Morel	No, this is a False Morel
Medium	Explanation of correct answer	Explains why the correct answer is correct.	No, this is a False Morel. False Morels have asymmetric caps that bulge to one side and the cap is not attached to the stem. It has wavy fissures like a brain.	No, this is a False Morel. False Morels have asymmetric caps that bulge to one side and the cap is not attached to the stem. It has wavy fissures like a brain.
High	Error-Sensitive Feedback	Attempts to correct specific error committed; compares and contrasts correct and chosen response.	This is not a True Morel, it is a False Morel. (Display pictures of true and false morels). Notice that True Morels have a cap that is symmetrical and longer than the stem. In contrast, the cap of a false morel bulges to one side and is not attached to the stem.	This is not a Coral Fungus, it is a False Morel. (Display pictures of coral fungus and false morel). Notice that the coral fungus has branchlike angular spines. The false morel has wavy fissures like a brain.

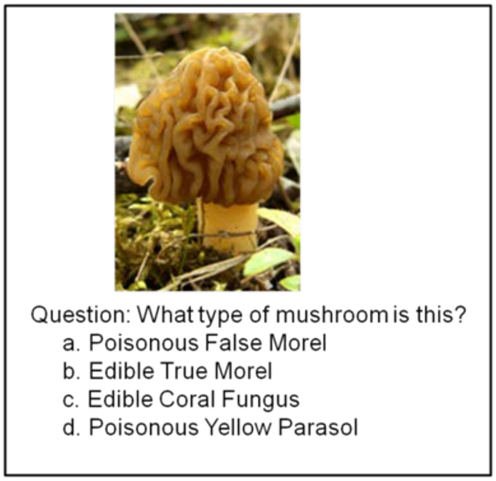

Figure 4. Example of a multiple-choice question. The correct answer is a Poisonous False Morel. See Table 1 for potential methods of giving feedback to erroneous responses.

The forms of explicit feedback listed in Table 1 are "locally-adaptive." They are *adaptive* because the feedback provided depends on an evaluation of the student response. They are *local* because the feedback given takes into account student data from a single point in time (Durlach & Ray, 2011). For an answer-based tutoring system this response could be the student's final answer; for a step-based tutoring system this could be their input on a step. Locally-adaptive feedback can take into account more than accuracy (e.g., it could also consider reaction time or certainty). What makes it local is that the decision regarding what feedback to present uses data collected from one point in time. This is usually the most recent student input. It does not consider any historical, previously stored information about the student or student-system interactions. Empirical research suggests that error-sensitive feedback is the most effective form of locally-adaptive feedback (Durlach & Ray, 2011). Although it takes more effort and resources to create (because different ways of erring must be followed up with different feedback messages), the effort is rewarded with improved learning.

In contrast to locally-adaptive feedback, a feedback message could consider stored information derived from previous student-system interactions or other sources. Such stored information is sometimes referred to as a student model (Woolf, 2009). A student model typically aims to keep a record of the evolving state of student mastery over a set of learning objectives; although, it may also incorporate other types of data, such as student traits, interests, and motivations (Vandewaetere, Desmet, & Clarebout, 2011). We will refer to adaptive decisions that use the amassed data in a student model as "context-aware." With context-aware feedback, the student may be given different feedback for the same local input, under different circumstances. For example, if the student model suggests that a student already can identify false morels, then an error on the question posed in Figure 3 may produce feedback like, "hmm, I thought you knew this one." Or if the student model suggests the student can identify True Morels and Coral Fungus the system may provide feedback like, "You've already correctly classified True Morels and Coral Fungus. Recall that True Morels have a cap that is symmetrical and longer than the stem and Coral Fungus has branchlike angular spines." In contrast, the cap of a false morel bulges to one side and is not attached to the stem. The false morel has wavy fissures like a brain. Messages like these allow the instructional system to more closely replicate

the feedback that a human tutor might provide, and to personalize the experience. As another example, context-aware feedback might affect feedback timing: if the student is nearing a mastery criterion, then the feedback might be delayed until the end of a scenario-based training, whereas a more novice student might be provided feedback upon each decision relevant to the learning objectives. A system that supports context-aware feedback will be more complex and resource intensive to create, because additional pedagogical rules will have to be implemented to determine how to use the contextual information to adapt the feedback. At this point in time, empirical evidence to justify this extra effort is lacking (Durlach & Ray, 2011). Therefore, our framework does not attempt to delineate all the possible ways context-aware feedback could be delivered. Research is required to determine the circumstances and rules under which context-based feedback is rewarded with enhanced learning outcomes.

The five levels of corrective feedback specified in FIT are listed in Table 2. To summarize the foregoing discussion, there is empirical evidence suggesting that Level III enhances learning relative to Levels 0, I, and II; but there is a lack of evidence as to whether Level IV provides additional benefits commensurate with the extra design and development requirements for its implementation.

Table 2. ***The Five Levels of Corrective Feedback Specified in FIT***

Level 0	No explicit item-level feedback – only summary score
Level I	Minimal feedback (item accuracy information)
Level II	Correct answer or explanation of correct answer
Level III	Error-sensitive feedback
Level IV	Context-aware feedback

Feedback in the FIT concentrates on corrective feedback as a form of adaptation. Positive feedback (confirmation that a response is correct) is not explicitly addressed because usually it is not adaptive. All students that respond correctly get the same positive feedback. It should be noted, however, that the content of positive feedback could be adapted based on some assessment of student certainty. The degree of confidence that students have in the correctness of a response can affect receptivity to feedback content (Kulhavy & Stock, 1989). If confidence is high and the response is correct, the student might pay little attention to the feedback. Therefore, there seems little point in providing more than minimal feedback. On the other hand, if confidence is low, students may benefit from explanatory information (why the response was correct). There is research to suggest that adapting the amount of explanation in positive feedback messages to the certainty of the student can be beneficial (e.g., Forbes-Riley & Litman, 2011; Mitrovic, Ohlsson & Barrow, 2013). Mitrovic et al. (2013) discuss several ways that an instructional system can be implemented so as to detect and respond to uncertainty.

Support

In ideal one-on-one tutoring, the tutor acts as a knowledgeable learning partner, who adjusts the learning activity according to the comprehension of the learner. The activity should require some mental effort from the learner, yet should be one that the learner can master, with support and guidance from the tutor. The approach of adapting the learning activity to being just

one step beyond the learner's current abilities is known as the mastery approach; and the support and guidance offered by the tutor is known as scaffolding. Scaffolding can encompass a number of different instructional tactics, such as diagrams, examples, attention focusing, hints, pumps, and encouragement. Effective scaffolding reduces confusion and uncertainty, and keeps the student on task. It can also encourage the student to engage in metacognitive processes important for learning, such as elaboration, self-checking, and self-reflection (Aleven, Stahl, Schworm, Fischer, Wallace, 2003; Azevedo & Hadwin, 2005; Luckin & du Boulay, 1999; Renkl, Hilbert, & Schworm, 2009).

The educational literature prescribes that as learner competency increases, scaffolding should gradually be withdrawn (Pea, 2004; Wood & Wood, 1999). This is often referred to as fading. Just as a person healing from a broken leg may go from crutches to a cane to no assistance, ultimately, the learner should be able to apply their knowledge to solving a problem without any assistance. Thus, by its very nature, scaffolding should be context-aware—increased when the student is having difficulty and decreased as their mastery increases. Deciding the level of support thus depends on knowledge of a student's evolving competence. In technology-based instruction, support can be provided on request, upon a partial correct solution, or proactively, e.g., if the student takes too long to make the next response (e.g., Fossati, 2009). In several step-based ITS, students can request hints multiple times on each step. Repeated selection of the hint button provides an increasingly directive hint. The final hint—the bottom-out hint-- provides the solution if necessary (e.g., Guo, Heffernan, & Beck, 2008; Roll, Aleven, McLaren & Koedinger, 2011).

A different ITS approach for hinting has been to associate decreasingly detailed hints with problem number; i.e., most detailed hint on problem one, less detailed hint on problem two, etc. (e.g., Schwonke, Renkl, Salden, Aleven, 2011). This is meant to be analogues to fading (reducing support as the student learns). While these types of support mechanisms are sometimes referred to as scaffolding; they are not true scaffolding, because they do not depend on the student's current level of understanding. They are not adaptive, since all students have the same sequence or detail of hints available. Wood and Wood (1999) did implement adaptive fading in their ITS, QUADRATIC. The level of detail provided in a requested hint was contingent on the student's performance and the level of hint they received on the previous exercise. Problem success meant the next requested hint was given at a less detailed level than the hint on the last problem, whereas lack of success meant the next requested hint would occur at a more detailed level than before.

To make the distinction between different implementations of support clear, our framework explicitly separates methods like access to fixed hints vs. truly adaptive scaffolding techniques. We also make the distinction between locally-adaptive and context-aware adaptive support. A locally-adaptive hint or prompt is contingent on the latest student response (or lack of a response). Wood & Wood's QUADRATIC ITS is an example, because it uses the student's response on the last problem (as well as knowledge of the detail level of the hint given). A context-aware hint or prompt is contingent on the pattern of student performance over time. For example, QUADRATIC's implementation might be modified to provide hints on the basis of a student model of past performance and hinting detail (instead of just the last problem). For complex tasks, which draw on knowledge related to multiple learning objectives, different levels

of support might be assigned to different exercise subcomponents. Steps related to already mastered learning objectives may have little support provided, whereas those related to less secure understanding may have relatively more support provided. Analogous to Context-aware adaptive feedback (Level IV), context-aware adaptive support enables the system to interact in a more flexible way with the student, capitalizing on multiple prior interactions to construct hints and prompts. When such a system is equipped with a natural language interface, this method of scaffolding can be implemented as interactive dialogs (in text or orally). An example of a system that uses interactive dialog is the ITS, AUTOTUTOR, described in Graesser, Jeon, and Duffy (2008). Dialog between an ITS agent and the student supports the student in refining and explaining problem solutions.

Sometimes feedback and scaffolding are grouped together as support techniques (VanLehn 2006; 2011). Support tends to be given during problem solving, whereas feedback can be given either during or after problem solving. The distinction can be blurry, however. An After Action Review (AAR), which helps a student critically analyze what they did right and what they did wrong, tends to mix feedback and support. While acknowledging that support and feedback can be integrated, our framework separates the two, because some instructional systems may include feedback only, and we want our framework to be able to clearly characterize instructional functionality.

The five levels of support specified in FIT are summarized in Table 3. Level 0 represents no support. Level I represents support that is pre-scripted and accessed on the student's initiative. An example might be a glossary, a link to additional explanatory information, or a "request a hint" button. In the latter cases, information accessed through the link or button is problem-centric. It is intended to assist students with the problem at hand, but is not adaptive, because the information or help available is the same for every student. It is adaptive only in the sense that the student has the option to use it or not. Level II represents resources that are locally-adaptive. The support available depends on some aspect of the student's most recent task performance. The support can be either requested by the student (IIa) or triggered automatically (IIb), or both. The ITS Quadratic, described above, is an example of IIa. Latency-triggered reminders would be an example of Level IIb; e.g., in an emergency response scenario-based exercise, a first-responder may be reminded to call in a status report, if they neglect to do so after a criterion amount of time. Level III support is responsive both to local and contextual information. It is true scaffolding, in which the student is weaned off support as mastery increases across practice. A sailor learning to navigate by magnetic compass for the first time may be guided through the steps of applying local variation and deviation to calculate the required compass heading, whereas a student getting refresher training may simply be given a memory jog (e.g., a mnemonic for True, Variation, Magnetic, Deviation, Compass is *TV Makes Dull Children*). Finally, Level IV represents true scaffolding mediated by natural language dialog. This level of support allows mixed-initiative dialog between the student and a pedagogical agent, helping the student to actively construct knowledge (Graesser et al., 2008).

Table 3. ***The Five Levels of Support Specified in FIT***

Level 0	No support
Level I	Fixed hints on request (problem determined); other fixed sources of information (e.g., glossary) where student initiates access
Level II	Locally-adaptive hints, prompts, or pumps a. on request b. triggered
Level III	Context-aware adaptive hints, prompts or pumps (True Scaffolding) a. on request b. triggered
Level IV	The same as Level III, with interactive dialog

There is evidence that including some form of support in instructional technology is beneficial for learners (e.g., Aleven & Koedinger, 2002; Clarebout & Elen, 2009; Kali & Linn, 2008); however, few experiments have attempted to examine whether *adapting* support affects learning outcomes. One, conducted by Luckin and du Boulay (1999) had three hint conditions, which were analogous to FIT's Levels I, II, and III; but, their test was completed by only 26 students, and their results were inconclusive. Relatively more research has been completed in the area of "faded worked examples." Worked examples are step-by-step demonstrations of how to perform a task or solve a problem. Fading worked examples is an instructional technique in which the amount of the problem that the student has to solve is gradually increased. Over time, the student progresses from reviewing completely worked out examples to solving entire problems. The process can also include requiring students to justify solution components. When the fading is conducted according to a fixed schedule, it is not adaptive, whereas when the fading is governed by the student's own past performance on previous problems, it is adaptive. Salden, Aleven, Renkl and Schwonke (2009) demonstrated a learning benefit from adaptively fading worked examples in the context of students solving geometry problems requiring the application of four different theorems. In the adaptive condition, transition from presenting a solved problem step vs. requiring the student to solve the step was based on an estimate of whether the student understood the relevant theorem. That estimate was based on whether the student previously was able to choose the right justification (from a menu) for an analogous step in previous examples. Students in this condition performed better on a posttest than students who had received fading of worked examples according to a fixed schedule. Two other experiments have also shown learning benefits from adaptive fading of worked examples (Corbalan, Kester, & van Merriënboer, 2008; Kalyuga & Sweller, 2004). In these three experiments, the adaptive conditions could be labeled as FIT support Levels IIb (Kalyuga & Sweller, 2004) or IIIb (Corbalan et al., 2008; Salden et al., 2009); however, the fixed conditions do not really fit

naturally as support Level I, because they were not under student control. In the context of FIT, these results are more properly viewed as evidence for mastery learning. That is, they are more naturally interpreted in the context of FIT as adaptively increasing the challenge for the student as they show more evidence of competence, as opposed to reducing support (fading). Nevertheless, we mention the results here because of the superficial similarity between "fading of worked examples" and fading of scaffolding.

Rather than examining how adapting support affects learning, research has tended to focus more on how students use on-demand help (Level I, or IIa) and how the content of on-demand help affects learning outcomes. With regard to how students use on-demand help, it has been found that many students abuse the help, by rapidly requesting repeated hints until they get to the bottom-out hint. Thus, they overuse the hints to get through problems. Other students do not use help enough; i.e., they neglect to use hints when they might be of help (Aleven et al., 2003; Clarebout & Elen, 2009). To try to correct these misuses of hints in ITS, Roll et al. (2011) implemented a Help Tutor within a geometry ITS. This Help Tutor provided students with feedback on their use of the available hints, instruction on help-seeking, and support for self-assessment. This supplement to the original geometry ITS reduced help abuse; but it failed to have a significant effect on learning outcomes. With regard to research on designing the content of hints or help messages, content that fosters student reflection, self-explanation, and mental model formation tends to be most effective (e.g., Dutke & Reimer, 2000; Renkl et al., 2009; Roll et al., 2011). Dutke and Reimer's (2000) results suggest that the optimal design of on-demand help may differ depending on whether the help is intended as a performance aid only (to accomplish the task as quickly as possible) vs. whether the help is intended to instruct (so that future tasks can be accomplished without help). If the latter, help content should focus on principles and support construction of student understanding.

Micro-sequencing

As mentioned in the previous section, the mastery learning approach involves tuning the learning activities to the student's current capabilities. It is one adaptive technique supported by ample evidence as to its effectiveness (Guskey & Pigott, 1988; Kulik, Kulik, & Bangert-Drowns, 1990; Perrin, Dargue, & Banks, 2003). The implementation of the mastery technique during one-on-one tutoring is likely one of the main reasons it is so successful (VanLehn, 2011). The technique involves organizing learning objectives into sequential modules, which build on each other. Students are required to demonstrate mastery on one module before proceeding to the next, more advanced, module. Mastery is demonstrated through various forms of ongoing assessment. The assessments are designed to diagnose student strengths and weaknesses. After initial assessment, the next learning activity should be tailored to repair any diagnosed problems. These activities are called correctives or remediation. The aim is to select the best learning activity that will keep students at the right level of challenge, avoiding both boredom and frustration. It is assumed that with such targeted instruction, all students should be able to master a module and move on (Bloom, 1976). Mastery learning is relatively easy to implement in instructional technology for individuals, compared to classrooms, because each student can progress at his or her own pace. Effective use of the technique requires valid automated assessment measures of learning, however, which take skill and effort to create. These measures should be linked to the learning objectives and ideally have known psychometric properties

(such as average difficulty, ability to discriminate different levels of mastery, or strengths on component skills or knowledge). Many current on-line learning applications seemingly implement mastery learning (e.g., by requiring a certain score on a module posttest); however, they rarely include tailored remediation for students who do not meet the mastery criterion.

The FIT micro-sequencing construct addresses the different ways that instructional technology can respond to students who fail to demonstrate performance at a mastery criterion. Micro-sequencing can be concerned with remediation (repair) or enrichment (i.e., more in depth materials for high performing students); however, FIT is primarily concerned with methods of remediation. By definition micro-sequencing is adaptive. If all students receive exactly the same experience, no sequencing decisions are made during the learning experience.

Four main decisions are required to implement micro-sequencing. First, what constitutes mastery? Second, is remediation mandatory or optional? Third, are dedicated remediation (or enrichment) materials available? Fourth, how will mastery be re-assessed?

1. What constitutes mastery? The issue here is whether mastery is determined by a pattern of assessment scores, with different scores associated with different learning objectives, or whether mastery is represented in one aggregate score. Assessment can be conducted as a separate event (such as a test), or it can be integrated into the learning activities themselves. Interactive digital learning activities (e.g., step-based problem solving, simulation and micro-worlds), provide the possibility to assess students at the same time as they apply or practice what they have learned. If mastery of a Terminal Learning Objective (TLO) is achieved through demonstrated mastery of associated subtasks or Enabling Learning Objectives (ELOs), then mastery can be represented by a pattern of scores across ELOs. In that case, the module might be deemed mastered only when all of the ELOs have been mastered to the designated criteria (regardless of the average across them). Alternatively, the mastery criterion may be represented by a simple average across assessment items. A benefit of the pattern approach is that it allows diagnosis of student problems and subsequent customization of the next activity so as to focus on areas of student weakness. If mastery is represented as a single score, the information about specific strengths and weaknesses may be lost, and therefore the next activity cannot be customized to the same degree. The different levels of micro-sequencing in FIT roughly correspond to how finely tuned the next activity is to student needs.

The contents in Table 4 and Table 5 are used to better illustrate the benefits of using the pattern approach compared to the summative approach. Table 4 presents an example TLO, broken down into two ELOs, which are further broken down into knowledge components (i.e., the specific pieces of information that need to be learned). Table 5 presents hypothetical assessment results of two students who have completed a training simulation designed to assess their knowledge of the content. In the far left column, Table 5 lists 12 critical decisions points in the simulation scenario, with each decision point potentially demonstrating understanding of the learning objectives as laid out in the Table 5 column "Assessment related to." It can be seen that although the two students have the same average score across the decision points (75%) their patterns across the decision points are quite different. Student 1 has difficulty only with the most difficult decisions, requiring integration of two pieces of knowledge, whereas Student 2 has problems specifically with ELOs 2b and 2c. By understanding this pattern of results, the two

students can be given different interventions targeted at their specific problems. Such targeted interventions are not possible when only the overall average scores are considered. Table 5 also illustrates that the mapping of assessment to knowledge components need not be one-to-one. Complex mappings are possible (many-to-many). In addition, weights might be applied to represent the relationship among different knowledge components and assessments. Finally, assessment patterns might influence the selection of corrective remediation on the basis of raw scores or on the basis of modeled scores. For example, a modeled score may take the chance of guessing correctly into account, and/or the degree of relatedness among ELOs (i.e., the degree to which mastery of one ELO predicts understanding of another ELO). Some models aim to predict the next best learning activity for the student, given the pattern of their past responses and knowledge of remediation tasks, such as average task difficulty (e.g., desJardins, Ciavolino, Deloatch, & Feasley, 2011).

Table 4. ***An Illustration of Hierarchically Arranged Learning Objectives. With the Terminal Learning Objective (TLO) Broken Down Into Two Enabling Learning Objectives (ELOs). ELOs Are in Turn Broken Down into Smaller Knowledge Components***

TLO: Understand Airspace Procedural Controls	
ELO 1: Understand the three types of airspace separation	ELO 2: Understand five types of airspace control measures
1a. Temporal	2a. Air control point
1b. Lateral	2b. Air corridor
1c. Vertical	2c. Coordinating altitude
	2d. Restricted operating zone
	2e. Minimum risk corridor

Table 5. ***An Illustration of the Assessment Results of Two Hypothetical Students. Both Students Have the Same Average Score across Assessment Items (Decision Points); However, Their Patterns across Items Suggest They Would Benefit From Different Forms of Remediation***

Decision Point	Assessment related to:	Student 1	Student 2
1	1a	✓	✓
2	1b	✓	✓
3	1c	✓	✓
4	2a	✓	✓
5	2b	✓	X
6	2c	✓	X
7	2d	✓	✓
8	2e	✓	✓
9	2c & 1c	X	X
10	2d & 1b	X	✓
11	2b & 2e	X	X
12	2d & 1a	X	✓
Simple Average		75%	75%

<u>*2. Is remediation mandatory or optional?*</u> Having gone through some core content and assessment, some instructional applications provide an underperforming student the choice between going through remediation or conducting re-assessment immediately, without remedial review. The re-assessment may be confined to the learning objectives that were failed, rather than a re-assessment on the entire scope of the module. Allowing immediate re-assessment may be unproductive from a learning point of view, because it may encourage students to guess, especially when multiple attempts at re-assessment are allowable. This is particularly a problem when students select from a limited number of responses to assessment items (e.g., multiple choice). It is less of a problem when students are required to generate responses (such as producing steps to solve an algebra problem), as long as new assessment items are available.

<u>*3. Are dedicated remediation materials available?*</u> The simplest form of remediation is recycling. In recycling, students go through content over again. If mastery is based on a simple summary score, they may have to review the entire module, or at least navigate through the entire module. If mastery is based on a pattern of ELO scores, and the learning materials are constructed appropriately, it should be possible for students to review the materials specifically related to their areas of weakness. Recycling may be effective if mere forgetting is the source of sub-criterion assessment performance. On the other hand, it is unlikely to be effective if the source of sub-criterion assessment performance was due to an inability to understand the previous content. In that case, dedicated remedial materials may be more useful. These would consist of supplemental content, specifically designed for remediation. An example is an application created by Tseng, Chu, Hwang, and Tsai (2008) to help junior high school students learn about mathematical sequences. The system had three ways of presenting content: “Easy,”

which reviewed prerequisite concepts in detail while introducing each new concept at a basic level, “Middle,” which reviewed only the most relevant prerequisites, while introducing each new concept, and “Difficult,” which did not review prerequisites and presented both basic and advanced concepts. If a student failed a Difficult- or Middle-level module posttest, students repeated the module, but at the next lower level. Conversely, students who did very well on a module posttest would get the next module at a higher level. Students learning with this system outperformed those learning with the system fixed at the Middle level throughout (by about 0.8 standard deviations), as measured by a post-learning test. Another approach to creating supplemental remediation is to create different content versions using different types of media. For example, one version may explain a concept in text, while another version may present a visual demonstration, and yet a third version may provide a practice environment. Students may opt to utilize any of the different versions on their path to mastery, and be offered one not yet utilized for remediation. Finally, supplemental remediation may simply involve more problems, in contexts where students must solve problems or perform a task (i.e., students are given problems or tasks until they reach mastery).

4. How will mastery be re-assessed? The issue here is whether the same exact assessment measures are used to assess mastery post-remediation as compared with pre-remediation. Parallel forms of assessment are desirable (i.e., the same underlying competencies assessed by different specific questions or methods). Some instructional technologies make use of test banks for re-assessment, with previously used questions excluded. It may be challenging to create parallel and equally difficult items; but it is especially important if remediation is optional and immediate retest is allowable.

FIT does not attempt to represent all the possible combinations of these four factors, but instead focuses on the nature of the remediation experience *per se* (question 3 above). Given that the student has been assessed as below threshold, what are the various ways that correctives can be given? The way correctives are implemented is really the essence of the adaptation. Various possibilities are listed in Table 6. The most basic form of adaptive micro-sequencing (Level 0) is Recycling. As illustrated in Figure 5a, when Recycling is implemented, students who fail to reach mastery simply repeat the content again. Recycling is usually based on a single aggregate performance score; however, a variation could be based on a pattern of ELO scores, in which case the repeated content might be limited to that associated with the failed ELOs. In practice, Recycling is the only FIT level that commonly allows immediate re-assessment (dashed lines in Figure 5a) without required review. If the developers have gone to the trouble of creating alternative versions of the content, they typically require students to use it.

Level I, Supplemental Remediation (Figure 5b), is somewhat more adaptive than recycling. Instead of merely repeating content, supplemental materials are provided. The supplemental materials could be designed specifically for students who have been unable to master the learning objectives based on the core content of the module. For example, they may use simpler language, or more explanation. Alternately, they could use a different format than the core content, or just present additional opportunities to practice a task (e.g., more multiplication problems to practice). Level II, Supplemental Remediation Levels is even more adaptive (Figure 5c). It is similar to Supplemental Remediation, except that multiple versions of supplemental remediating materials are available. The version selected may depend on the

distance the student is from mastery, or on whether the student has already gone through one of the other versions. The Tseng et al. (2008) system described above is an example of Level II, because there were two levels of remediation available.

Table 6. ***The Five Levels of Micro-Sequencing Specified In FIT***

Level 0	Recycling
Level I	Supplemental Remediation
Level II	Supplemental Remediation Levels
Level III	Adaptive Content
Level IV	Real-time Adaptation

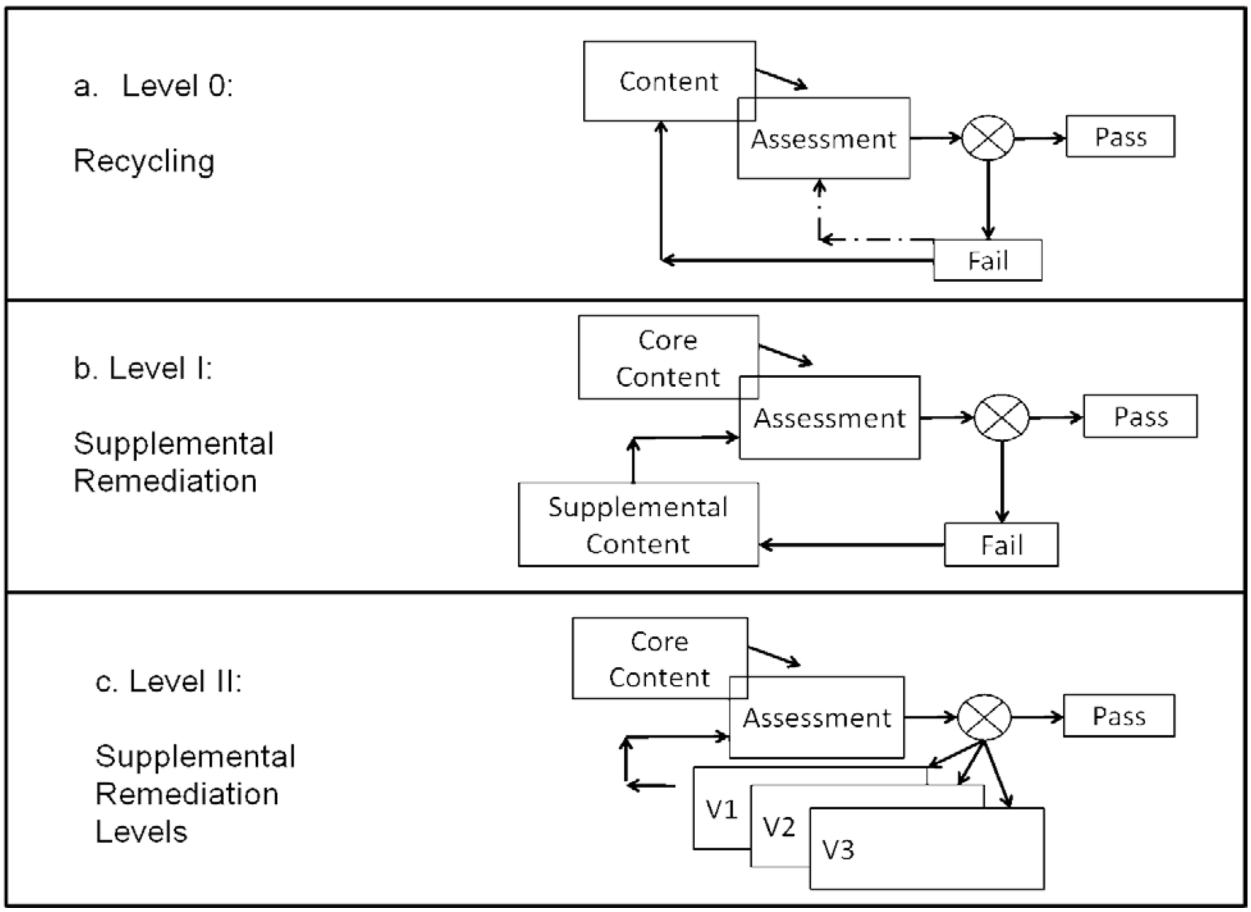

Figure 5. Levels 0, I, and II of micro-sequencing in panels a, b, and c, respectively.

As already described, the more granular the pattern of assessment scores, the more personalized the learning path can be, because more information is known about the student; however, personalizing the learning path also depends on opportunities to adjust the learning content or activities. If content or activities are formulated in blocks, the sequence can only be adapted between blocks. Therefore, the finer grained the blocks, the more opportunity there is to adapt. Levels III and IV, Adaptive Content and Real-time Adaptation, attempt to represent

micro-sequencing with finer grained blocks than Levels 0 – II. They both require that learning activities be associated with knowledge components, similar to the example given in Table 5. They are distinguished by whether “blocks” are discrete or continuous.

With Adaptive Content, each student experiences his or her own sequence of problem-based tasks. These tasks have integrated assessment such that assessment results on prior tasks are used to select the next task, until the student reaches mastery. The ultimate mastery goals are the same for all students, but the path by which they get there may be different. So for example, in a geometry ITS, each student would receive a different sequence of practice problems, depending on the types of errors they commit. Corbett and Anderson (1995) refer to this type task selection in an ITS as “cognitive mastery learning.” Some ITS accomplish this with the aid of a “bug library” (e.g., VanLehn, 1982). Bug libraries represent common student misconceptions or confusions. They enable ITS to interpret not only that an error occurred, but also the underlying misconception that caused the error. When the system can recognize an erroneous response as indicative of a particular misconception, it can provide remediation targeted precisely to fix it.

Figure 6 illustrates Adaptive Content micro-sequencing in the context of a shoot-no shoot decision exercise. Imagine that immersive practice on this decision making task can be varied parametrically in terms of number of targets and number of distracters, and that there is a library of 16 exercises represented by combining four levels on each of these factors. Depending on how well students perform at selecting targets and avoiding distracters, the number of targets and distracters will be increased (or not) from one practice exercise to the next. Different students take different paths through the space of possible scenarios with two possible paths illustrated in the figure. An application of Adaptive Content such as this can be found in Levchuk, Shebilske, and Freeman (2012). They applied modeling techniques to determine paths for team training scenarios for the Air Force Dynamic Targeting Cell, where the scenarios varied planned targets requiring offensive action, and unplanned targets, requiring defensive action.

Real-Time Adaptation is similar to Adaptive Content, except that adjustment of the learning task is transparent to the student. If Real-Time Adaptation applied to a shoot-no shoot decision exercise, the challenge level would adjust in real-time, without having to complete one exercise and start another. With a more narrative-based scenario, Real-Time Adaptation would involve adjusting scenario events “on the fly,” based on the student’s ongoing performance. For, example, in a first-person shooter type of training scenario, the skill level of the opposition forces might be incremented or decremented according to the student’s success thus far. As another example, in a mission command scenario, the complexities of managing infrastructure rebuilding and host-nation institutions may be amped up gradually as the student demonstrates the ability to handle simpler situations. Accomplishing scenario-based training with Real-Time Adaptation is currently a matter for research (e.g., Domeshek, Durlach, & Bratt, 2010).

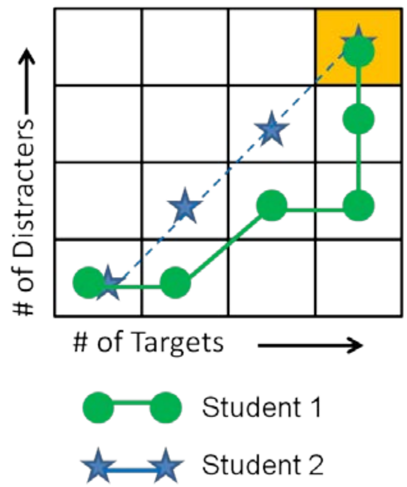

Figure 6. Example of how micro-sequencing can produce two different paths in shoot/no shoot decision training. Both students start out at the easiest level (lowest number of targets and distracters, represented in the lower left corner). Student 1 has more difficulty withholding fire at distracters than Student 2, so is advanced more slowly up the distracter challenge levels

There are multiple important considerations in deciding what level of micro-sequencing to use. First, in general, the more adaptive, the more content required. Anything more advanced than recycling requires that "extra" content be created. E.g., the shoot/no-shoot example in Figure 6 (Adaptive Content) requires 16 different scenarios, even though some students may only need to complete as few as four. Depending on the nature of the training, Real-Time Adaptation could be more efficient. In the shoot/no-shoot example, only one scenario might be required, if it can automatically tune the level of challenge to the student's current level of skill, and add a degree of randomness to where and when targets and distracters appear.

A second consideration in selecting a level of micro-sequencing is that, for Adaptive Content and Real-Time Adaptation, the designer must be able to specify the association between performance assessment measures and the knowledge components (as in Table 5). These associations are used to update the student model based on performance. Likewise, the designer must be able to specify the associations between knowledge components and content. These associations are used to determine which content would best serve to repair student weaknesses or which path the student should take through the learning space. To the extent flaws exist in these associations, the benefits of adaptive micro-sequencing will be undermined. Ensuring accuracy typically will require a more intense effort of up-front domain analysis (e.g., cognitive task analysis) and a more rigorous approach to assessment design, compared with less adaptive methods. Data mining of performance results from training with models crafted by experts may subsequently be applied to improve these associations (e.g., Koedinger, McLaughlin, & Stamper, 2012).

Another consideration is that any technology-based instructional system that requires mastery for a student to progress through the modules results in students finishing at different

times. If students are required to engage in some follow-on group activity, dependent on the knowledge gained, that activity needs to be scheduled far enough out so that all students have sufficient time to complete the course of instruction.

As mentioned at the beginning of this section, there is good evidence on the effectiveness of the mastery approach (Guskey & Pigott, 1988; Kulik et al., 1990; Perrin et al., 2003). Yet there is surprisingly sparse evidence when it comes to comparing the various methods of implementing it. Few studies have endeavored to compare the relative effects of the different FIT levels of micro-sequencing, holding all other factors constant. As already mentioned, Tseng et al. (2008) compared the effects of presenting students with fixed content (Middle) vs. adapting the difficulty of the content (Easy, Medium, or Hard) according to student performance, and did find a large effect size (0.8). They applied Level III, Supplemental Remediation Levels micro-sequencing for determining the remediation content if a student failed a post-unit test, as well as for selecting the difficulty level of the content presented in the next unit, once the previous unit was passed. They compared this with Level I (Recycling), in which all materials were presented at the same (Middle) level. Because they adapted both remediation and new content selection, it is unclear what the relative contributions of the two adaptations were compared to simple recycling. I.e., we don't know whether adapting remediation content, but presenting all new content at the Middle level would be just as effective.

Given the amount of effort dedicated to creating ITS, which implement Level III Adaptive Content micro-sequencing, there has been surprisingly little empirical test of whether customizing the sequence of problems students work on produces any benefit over simply imposing a mastery approach, and presenting problems from a problem bank (i.e., micro-sequencing Level I or II). The majority of empirical tests have compared the efficacy of ITS to no intervention (e.g., ITS added to standard classroom teaching vs. classroom teaching only) or to one-size-fits-all, non-mastery based methods (for a review see Durlach & Ray, 2011). After conducting a meta-analysis of various tutoring methods, VanLehn (2011) concluded that it is the granularity of the feedback and support that seems most responsible for enhanced learning outcomes, although he did not explicitly attempt to isolate content selection (micro-sequencing) as a factor. Thus, based on the lack of evidence, it is unclear whether learning gains are to be had from micro-sequencing Level III over Levels I or II.

Macro-sequencing

Brusilovsky and Vassileva (2003) discuss the idea of dynamic courseware generation, in which a course could be generated "on the fly" to take into account the students' existing knowledge, goals, and timeframe, adapting dynamically to their difficulties and rate of progress. Their discussion of dynamic sequencing addressed both adaptations for dealing with remediation and adaptations for sequencing topics. Our intent is to separate the process that recommends what to do next when a student needs remediation (micro-sequencing) and the process that recommends what to do next once current topic mastery has been attained. We call that latter process macro-sequencing. Macro-sequencing has to do with determining the order in which new topics or learning objectives are introduced. Macro-sequencing can occur at different levels of granularity. It may involve curriculum sequencing (e.g., the path to completing a major in psychology), or the order of topic assignments within a course. Or it may involve decisions

about lesson and exercise sequencing within a module. Traditionally, topic sequences have been based on domain analysis and instructors' past experiences of what works well. It has been suggested that topic sequences should also be based on how students actually learn. Mirroring stages of understanding and sophistication, succeeding topics should represent levels of achievement that are intermediate steps in the pathway to expertise (Corcoran, Mosher, & Rogat, 2009). Our review of macro-sequencing techniques is not concerned with identifying the ideal learning progression through a specific domain, but rather, once those learning progressions are determined, how they might be adapted for different students, depending on their learning goals and learning performance.

Just as there are several questions that need to be considered when implementing micro-sequencing, there are also multiple questions involved in macro-sequencing. One of those questions concerns the extent to which students are allowed to choose the sequencing of topics. Like choosing courses to take for a university major, some topics may be mandatory, whereas others may be a matter of choice (with respect to whether to take it or when to take it). Certain pre-requisites may be mandatory before an optional advanced topic can be selected. In some cases, students may be given recommendations on the sequence of topics to follow, but still retain the choice of whether to follow the recommendations or not (e.g., Triantafillou, Pomportsis, Demetriadis, & Georgiadou, 2004; Tsiriga & Virvou, 2004).

Another question to be considered is whether multiple versions of the same topic will be available. Different versions may be created with the intention of optimizing the fit between student aptitudes and instructional strategy, media, and/or cognitive demands. An aptitude in this context is an individual characteristic such as cognitive or learning style, self-efficacy, motivation, prior experience, or cognitive ability. A few reasons to produce different topic versions are: to provide supplementary instruction to learners who are deficient in a particular aptitude, to provide a format that matches a learner's preferred mode of processing or perceiving (e.g., visual vs. verbal), and/or to challenge or stimulate learners with above average aptitude.

Multiple versions might also be created to suit students with different reasons for learning. For example, a student may need to learn a foreign language for an imminent job requirement or for an upcoming vacation. Depending on expected needs, students may cover different vocabulary and practice exercises. There may be some core modules completed by all students, after which students with different goals or roles branch to tailored materials. An example of this is Virtual Cultural Awareness Trainer (VCAT; Johnson, Friedland, Schrider, Valente, & Sheridan, 2011), which provides cultural and language training that focuses on the development of mission-relevant intercultural competence. Students receive practice in scenarios that are selected based on their expected areas of responsibility. Trainees starting the course complete a placement questionnaire in which they indicate their particular area of responsibility, mission focus, and level of seniority. The VCAT course delivery system then automatically selects the set of required modules that target language learning for contexts most related to the trainee's anticipated mission role.

If multiple versions of the same topic are created, then the question arises as to how information about student aptitudes or goals will be collected and utilized to recommend or mandate one of the available versions. One approach is to collect student information prior to

the instructional experience, as in the VCAT described in the previous paragraph. An alternate approach is to collect student data during the instructional experience and dynamically select topic versions based on that information. For example, in the already mentioned work of Tseng et al. (2008), performance on each post-unit test was used to determine whether the next unit was presented at the Easy, Middle, or Advanced Level.

FIT's methods of macro-sequencing intend to capture these various possibilities. The FIT macro-sequencing levels are listed in Table 7. Level 0 represents the case in which the order of modules is mandated, and all students follow the same sequence. Level I refers to cases in which there is some degree of, or complete student control over the order in which modules are presented. In the hybrid case, certain modules may be required before access to others, but there is still some element of student choice in the sequence. For both Levels 0 and I, all students go through all the same content, but perhaps in different order at Level I.

Table 7. ***The Five Levels of Macro-Sequencing Specified in FIT***

Level 0	Fixed sequence (one version only)
Level I	Student choice or hybrid choice/fixed
Level II	Test-out
Level III	Adapted Ahead (multiple versions)
Level IV	Adapted During (multiple versions)

Level II, Test-out, characterizes situations in which the student is given a knowledge pretest. Sets of pretest questions are associated with different modules, and a passing grade on a pretest set allows the student to skip the associated module. The aim is to save students time by allowing them to skip material they already know. When using Test-out, care should be taken to avoid the possibility of testing-out by chance (e.g., lucky guessing on multiple choice questions), and that the pretest assessment covers the learning objectives. One possible method of identifying lucky "test-out" is to re-test all the module content again as a post-test.

Levels III and IV both deal with cases in which multiple versions of instruction exist for the same topic, and information about the student is used to select which version the student experiences. Level III, Adapted Ahead, refers to situations in which pre-task measures or historical information about the student is collected before the instructional experience begins. The data are then used to select the sequence of versions. Once the version is selected, the content remains set throughout the training; no real-time adaptations exist based on student performance. Thus, Level III is analogous to Park and Lee's (2004) macro-adaptive instruction. Level IV, Adapted During, refers to situations where version selection occurs dynamically during learning, using data collected during student-content interaction. The data typically are about student performance during learning assessment. Students in the upper percentiles of performance may be given subsequent modules at more advanced levels, whereas students in the lower percentiles of performance may be given subsequent modules that have been simplified, contain more practice, or include other means to support learning.

These levels of macro-sequencing are not mutually exclusive. They could be combined in various ways. For example, Tseng et al. (2008) used Adapted Ahead on the basis of cognitive style and also Adapted During on the basis of module post-tests. As another variation, it is possible to use Test-out, but allow students to do modules in any order they choose. This combination was applied in an Army Civilian Education Course the authors were required to complete.

Research on the effectiveness of the different macro-sequencing approaches is sparse, except for Adapted Head (FIT Level III). Most of the evidence on this approach comes from the aptitude-treatment interaction (ATI) literature which seeks to determine whether different instructional formats (the treatment) should be selected for students of different aptitude, be it learning style, cognitive style, or cognitive ability as measured by past performance (Park & Lee, 2004; Pashler, McDaniel, Rohrer, & Bjork, 2008). With regard to there being learning benefits from adapting content on the basis of learning style, Pashler et al. (2008) described the evidence as unconvincing, despite the popularity of the idea, noting results that flatly contradicted the presumed benefits. With regard to adapting on the basis of cognitive ability as measured by past performance, Pashler et al. noted some evidence that low performers tended to do better in highly structured learning environments (one that provides explicit instructions and guidance) compared to less structured learning environments (one that provides little guidance), whereas the reverse was true for high performers. However, this was not universally the case and high performers tended to remain high performers regardless of the instructional strategy. Finally, with respect to cognitive style, Pashler et al. concluded there is modest evidence that adapting based on personality style produces learning benefits. However, the precise instructional conditions under which this occurs is not entirely clear. Overall, the ATI evidence involves a mix of findings, coming from research that varies on multiple dimensions (e.g., domain, instructional strategies, measures of aptitude, and outcome measures).

Far less research has been done on Adapting During, thus it benefits are unclear. As already mentioned, Tseng et al. (2008) found that adjusting topic difficulty up or down according to prior module performance led to better learning outcomes than keeping difficulty level constant at the middle level; however, they combined this technique with Micro-sequencing Level II, Supplemental Remediation Levels. The comparison condition did not receive this treatment so it is impossible to determine the relative contribution of the two adaptive procedures to learning outcome improvement. In general, adapting the level of challenge to student ability based on past performance is aligned with educational theory (Sweller, 1988; Vygotsky, 1978). Adapting During, by increasing challenge levels for high performing students may enable them to reach a higher level of expertise or sophistication, compared to not adapting. On the other hand without adapting, these high performers may complete the content in a faster amount of time. Therefore, the benefits of adapting for high performers depend on the goals: speed vs. level of attainment. Adapting During (by decreasing challenge levels for low performing student) may enable them to master a topic without remediation, which may improve motivation and speed to complete the course overall.

Sequencing vs. Branching

Readers familiar with designing IMI (or taking IMI courses) will notice that FIT does not attempt to exhaustively describe IMI varieties. Rather, FIT is specifically concerned with the *adaptive* instructional decisions that could be made when designing IMI. Nevertheless, in order to avoid confusion, this section highlights certain IMI sequencing decisions that bear superficial resemblance to micro-sequencing and explains how they are different. We will consider two examples in particular. The first concerns student choice in the sequence of content within a lesson and the second concerns branching scenarios.

Student control of sequencing within a lesson or module. Figure 7 illustrates a form of IMI that the authors have encountered frequently during required training. The figure is meant to represent a situation in which the student is directed to select a widget on the current screen to branch off and learn about a particular subtopic or to do a particular activity (e.g., find the information assurance problems in a simulated office). The student can select the widgets (A – D) in any order. In some versions, selection of A - D may be optional. In other versions, selection of each might be required before progression to the next topic is allowed. This type of sequencing is not a form of FIT micro-sequencing, because it has nothing to do with remediation. Rather, it provides the student a degree of choice in how they navigate the content. When choice of A – D is optional, this type of sequencing is most similar to Level I Support (fixed sources of information where student initiates access). When choice of A – D is required, it is most akin to Level I Macro-sequencing (student choice or hybrid choice/fixed), but occurring within a module. As already described, macro-sequencing can occur at different levels of granularity, and we therefore see this example as macro-sequencing within a lesson.

Branching scenarios. Another form of branching within a lesson is the so-called branching scenario, often associated with Level 3 IMI (TRADOC, 2010). Branching scenarios require students to apply knowledge in the context of a mission, an unfolding story, or problem. The student is presented with an initial situation, and then is required to make a decision about some action from a set of options. Depending on the choice selected, what happens next will be different. For example in one Army branching scenario for training unit leaders, *Danger Close*, the student takes on the role of a new platoon leader, and makes decisions about how to interact with his troops and his platoon sergeant, while preparing for and subsequently going on deployment. If the student makes poor choices, the platoon suffers a number of serious adverse outcomes during the deployment and even after its return due to manageable problems that were not handled effectively. At certain points, students can go back and review their decisions and receive feedback about factors to consider for each situation. This training system uses high quality video to immerse the student in the story; however, branching scenarios need not have such sophisticated media. An explanation of each situation, the decision to be made, and the options to choose from are all that is required. Much of the art in creating these scenarios is in deciding a coherent branching structure that is manageable. Figure 8 illustrates one possible branching structure, and how it can rapidly become complex even with only three choices per situation.

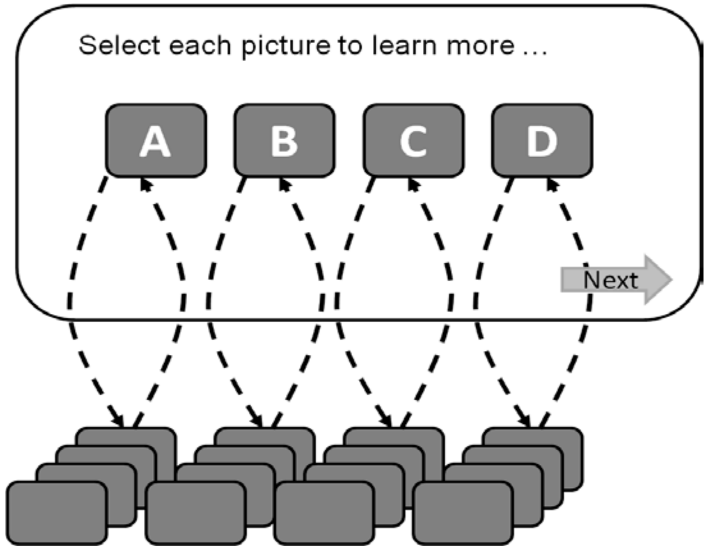

Figure 7. Schematic illustration of student control over sequencing during a lesson. Selection of A, B, C, or D leads to more content. The student is returned to the choice screen upon completion of each branch. Students can choose A, B, C, or D in any order. If selection of each is mandatory, the "NEXT" arrow will be disabled until all have been chosen.

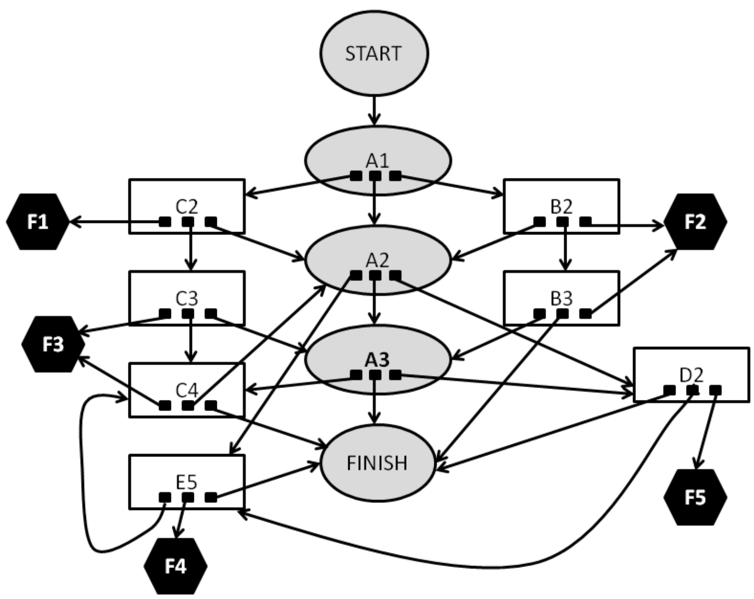

Figure 8. Schematic illustration of a notional branching scenario with three options at each decision point. The story path depends on which option (black square) is selected at each situation. Students can reach a successful conclusion in as little as three decisions (the A-path); but also have several routes to failure (F's).

The sequence of situations that a student encounters in a branching scenario is not a form of adaptation, as characterized by FIT. This is because the student's actions affect the sequence of events. Each situation that arises from the previous decision may provide implicit feedback

about whether the prior decision was good or poor by the very nature of the ensuing situation; however, that feedback is implicit (it does not provide explicit information about the goodness of the decision). Likewise the next situation is not a form of remediation, so this is not micro-sequencing. The point being made is that a branching scenario, in and of itself, is not a type of instructional adaptation, but rather a form of interactivity. The branching scenario can be combined with adaptive corrective feedback and support to promote learning.

THE RELATION BETWEEN FIT AND IMI

FIT is not meant to be a replacement for the IMI Levels rubric (TRADOC, 2010). Rather, it is meant to make more explicit the ways that instructional adaptation can be implemented in IMI applications. In this section, each IMI level is examined, and considerations for selecting FIT levels for feedback, support, micro-sequencing, and macro-sequencing are discussed. Here, we use the description of IMI Levels as found in TRADOC (2010).

IMI Level One – Passive

The learner acts solely as a receiver of information.

- Capable of computer generated multimedia presentations of intellectual skills (facts, rules, or procedures).
- Capable of showing a procedure with computer-generated multimedia explanations of equipment operations.
- Used primarily to introduce knowledge, including ideas, concepts, and processes.
- Information is generally provided in a linear format (one idea after another).
- Minimal interactivity is incorporated in the form of text, navigational icons, static graphics (e.g., photos, charts, tables) and illustrations, learner-initiated animations, and pop-ins and hyperlinks.

Assuming there is no assessment in Level One IMI, there is no opportunity to adapt on the basis of student performance. Therefore, feedback does not apply. Likewise there is no need for micro-sequencing, because there is no assessment of mastery. Various macro-sequencing approaches may apply. For example, students may be permitted control over the sequencing of modules (Level I macro-sequencing), or multiple versions may be designed for people in different roles or agencies (Level III macro-sequencing). The only type of support might be Level I (fixed sources of information), such as a glossary, or hyperlinks to more information.

IMI Level Two – Limited Participation

The learner recalls information and responds to instructional cues.

- Used to introduce simple operational and maintenance guidelines and procedures.
- Moderate interactivity is incorporated in the form of learner-initiated animations, interactive graphics, activities, scenarios, and assessments (practices, knowledge checks, and tests).
- Interactions force learners to make decisions related to material.

- This level has the capability of providing drill and practice, providing feedback on learner responses, emulating simple psychomotor performance, and emulating simple equipment operation in response to learner action.
- This level can be used for computer evaluation of intellectual skills using computer-based predictive and performance test items.
- Immediate or delayed feedback guides the learner to see the consequences and components of his/her performance.

Assessments are included in Level Two IMI, thus all aspects of FIT are applicable. The above description already alludes to the use of feedback, and its importance in guiding the learner to correct performance. We recommend FIT Level III feedback (error-sensitive feedback) as the best way of providing corrective feedback in the context of Level Two IMI. This helps the student understand not only what the correct answer or response was, but also why his or her answer was not correct. Especially for checks on learning during the course of the IMI session, they are most useful when students can learn from them to see where their knowledge is incomplete or flawed. Solutions requiring multiple steps should provide step-based feedback, not merely feedback on whether the final solution is correct or not. For example, in a "matching" exercise, where the student is asked to match terms and definitions, or match other types of items, it would be more helpful to illustrate which matches are correct or incorrect, rather than whether the whole pattern is correct or not. Also, showing the correct solution, without also displaying the student's answer, may make it difficult for the student to see where he or she went wrong.

Depending on the nature of the exercises or assessments the student is asked to complete, instructional support may be beneficial. It is likely most useful in the practice of multi-step procedures, or problems requiring a series of reasoned out steps (as opposed memorization, or learning of facts). For example, if you want a student to memorize that 3 + 3 = 6, support in addition to error-sensitive feedback might not be useful; however, if you want the student to understand *why* 3 + 3 = 6, then support can be used to facilitate the reasoning process (for example, by presenting graphical props). Therefore, a decision about whether to include support depends on the nature of the learning activities. Typically, support should be included to assist the student at arriving at an answer his or herself, and may depend on an understanding of what the student already knows to be effective.

Many Level 2 IMI applications require some criterion performance on a posttest (e.g., 80%) in order to receive certification. As previously mentioned, recycling (FIT micro-sequencing Level 0) may not be effective unless the source of error is mere forgetting. If lack of comprehension is the source of error, just repeating the same content again may not be sufficient to repair the error. Therefore, at least supplemental remediation (FIT micro-sequencing Level I) is recommended. Recall that for this technique, there is additional content available, which explains content in an alternate way or provides additional practice. Micro-sequencing Level II (supplemental remediation levels where there are multiple versions of supplemental materials) and Level III (adaptive content where the supplemental content is specifically tailored to the student's pattern of deficiency), are also effective methods of remediation; however, it is not clear whether they are sufficiently more effective than Level I to justify the extra time and resources required to implement them. Empirical evidence on this is simply too sparse to draw

such conclusions. Level III (adaptive content) is akin to what ITS tend to do when selecting problems for students to work on next (VanLehn's 2006 outer loop and as in the example in Tables 4 and 5). At issue is the return on investment for creating adaptive content vs. just providing more problems that cover the learning objectives without specific tailoring.

Level 2 IMI presents several opportunities for macro-sequencing. The Army Learning Concept 2015 (TRADOC, 2011) recommends tailoring learning to the individual learner's experience and competence level based on the results of a pretest or other assessment. Pretests can be used to allow learners to test-out of instruction they have already mastered, assuming that the pretest is valid and reliable (see TRADOC, 2011, page 21). Besides just allowing learners to skip content, however, knowledge about learner competence (and needs) can also be used to tailor content to their role or aptitude (macro-sequencing Level III, Adapted Ahead). If multiple versions of the same content are available (e.g., Basic, Intermediate, Advanced), and the course consists of multiple modules, both selection of the remediation version (micro-sequencing Level II, supplemental remediation levels) and selection of the level of the next module (macro-sequencing, Level IV, Adapted During) can be determined on the basis of posttest performance. Whether to use test-out vs. higher levels of macro-sequencing really depends on the instructional goals. The main purpose of test-out is to save time, whereas the main purpose of Adapting Ahead or During is to better match learner goals, interests, needs, or competence.

IMI Level Three – Complex Participation

The learner applies information to scenarios and interacts with simulations.

- This level is used to present more complex operational and maintenance procedures; also interpersonal interaction skills.
- Information is often non-linear.
- Moderate to high interactivity is incorporated in the form of complex interactive graphics including simulations and decision-based branched scenarios.
- Highly realistic scenario and equipment simulations fully involve the learner in near, part and whole task performance.
- After action feedback guides the learner to fully understand the consequences and components of adequate and inadequate performance.
- Feedback is based on tracking of several responses.
- This level is capable of providing complex branching paths based on learner selections and responses.
- This level is capable of evaluating learner intellectual skills and performance using computer-based performance and predictive test items.
- Computer evaluation of learner procedural performance includes the capability to generate time and error scores for performance test items.

All of the comments about feedback and support regarding Level 2 IMI also apply to Level 3 IMI. In particular, IMI Level 3 typically requires the learner to perform multi-step problem solving or decision making, and should be accompanied by step-based error-sensitive feedback. In the case of a branching scenario, this feedback should be delayed so as not to disrupt the narrative flow of the scenario, but not delayed too long, such that the student has difficulty remembering what ensued (Munro, Fehling, Towne, 1985; Salas, Burke, & Cannon-

Bowers, 2000; Schmidt & Bjork, 1992). The after action review should include the ability for the student to reflect on their prior decisions. It may also allow the student to "rewind" the scenario and choose different decisions.

Level 3 IMI typically involves assessing reasoning skills. Therefore, provision of support could be considered, but more research is required to understand the effectiveness of hints in the context of story-based scenarios. Level 3 IMI applications typically assume the learner has received some kind of didactic instruction already; but if not, then support may be especially useful. For example, the instructional application *BiLat* (Durlach, Wansbury, & Wilkinson, 2008; Kim, Hill, Durlach, Lane, Forbell, Core, Marsella, Pynadath, & Hart, 2009), a game-based simulation and tutoring system that trains cultural awareness and bilateral negotiation skills, depends on learners having been taught the principles of win-win negotiation. If they have not, the hints provided by the automated coach in that application can be particularly critical to a student making any progress in a negotiation scenario. FIT Level I support (fixed hints on request), might be designed to orient student decision making with regard to the important factors to consider (e.g., a hint may say "you should try to build rapport"). FIT Level II support (locally adaptive) could be applied by tailoring support based on the student's just prior input or decision (assuming it is relevant to the present one). For example, if the last decision was a poor one, then the hint might be more directive (e.g., "You should ask if he follows soccer"). *BiLat's* hints work in this way, becoming increasingly concrete if it appears that the student is floundering. Context aware support (FIT Levels III and IV) can be used, but is technically difficult to implement. It is not clear from evidence if these levels are justified by this extra effort.

Within the context of Level 3 IMI, micro-sequencing can be implemented in several ways. Recycling would constitute the opportunity for the student to do the same scenario again. This may be a more appropriate method of remediation in Level 3 IMI, than Level 2, because the student will experience a different branch of the scenario by making different decisions, compared to their first time through. Therefore in Level 3, recycling is not equivalent to mere repetition. *BiLat*, for example, uses recycling, in the sense that students have repeated meetings with the same simulated character until they meet their negotiation objectives with that character. *BiLat* remembers the outcome of the prior meeting, so that negotiation points already accomplished are retained, and any level of trust (or mistrust) built up with the simulated character from prior meetings is carried into the next scenario. Whether it makes sense to carry over the effects like this depends on the nature of the training.

Level 3 IMI applications represent an opportunity to put knowledge into practice, but students lacking the knowledge may struggle. One way of conveying that knowledge is through a worked example. In a worked example (or demonstration), the solution to a problem is displayed, with accompanying explanation of the rationale behind each decision or step. For *BiLat*, an introductory video was made, which not only explains the principles of win-win negotiation, but also presents negotiation scenarios: one that goes badly (with explanation as to why), and one that goes well (again with explanation). Depending on the student's prior knowledge, such worked examples can be used either as introduction or as remediation. An intermediate form is a worked example in which the student is asked to supply the rationale (e.g., from a menu perhaps). Students might be required to adequately explain why steps in a worked

example were correct or not, prior to taking a stab at solving a scenario themselves. If actual decision making is contingent on correctly explaining an example first, this would be an example of Level III micro-sequencing (adaptive content). Another version of Level III micro-sequencing would be selecting particular scenario dilemmas based on knowledge of the student's strengths and weaknesses in applying knowledge in the content domain. Of course, all this requires extra content over simple recycling, but little evidence exists regarding the payoff for the extra effort.

Any of the macro-sequencing levels of FIT can be applied to Level 3 IMI. As an example, *BiLat* uses student choice (Level II), in that the student can do the multiple scenarios available in any order they choose. An example of Level III (Adapted Ahead) would be if some scenarios were set in one cultural context (e.g., Iraq), whereas others were set in a different cultural context (e.g., Haiti). Soldiers might then do the practice negotiation most relevant to an expected deployment.

IMI Level Four – Real-time Participation

The learner engages in a life-like set of complex cues and responses.

- This level is used to simulate highly complex operational and maintenance procedures that often support certification.
- Maximum flexibility and multi-level branching allow a high degree of interactivity in the form of simulator and gaming environments.
- This level is capable of real-time simulation of performance in the operational setting and after action and natural consequences are given based on performance.
- This level incorporates artificial intelligence components and employs state-of-the-art technology for simulation and communication.
- This level can be used for computer evaluation of learner performance and intellectual skills using computer-based predictive and performance test items and the capability to generate time and error scores for performance test items.
- This level is often found in games with multiple players, computer-generated team players, and/or simulating decision-making incorporating multiple tasks.

All of the comments for Level 3 IMI apply to Level 4. Similar to Level 3, Level 4 assumes the learner already possesses content knowledge, and the exercise is an opportunity to put that knowledge into practice. Micro-sequencing can be used as remediation when the assumed knowledge appears to be insufficient, either through the use of worked examples with explanation, or through the presentation of supplemental didactic materials intended to reinforce the required knowledge. This remediation could be made adaptive (Level III micro-sequencing) by tailoring the remedial content to the specific knowledge components on which the student is the weakest. As in the example in Figure 6, real-time adaptation (micro-sequencing Level IV) can be applied to adjust the level of challenge of a scenario in real time, by making the scenario more or less challenging, as determined by student performance. This is akin to "leveling" in games, such that the level of challenge automatically increases once the player demonstrates proficiency at the current level. Different from typical games, however, the level of challenge might also decrease automatically, if the student demonstrates an inability to cope with the current level.

SUMMARY

When designing instructional technology, the designer needs to consider how learning effectiveness can be improved over one-size-fits all methods, by applying adaptive techniques that fit within the available time and resources. The primary objective for developing FIT is to illustrate, in a systematic fashion, the specific ways in which instruction can be made adaptive. While ITS are considered the epitome of adaptive training technology, it is important to realize that adaptive tactics exist that are easier and less resource-intensive to implement. Table 8 presents the typical adaptive behaviors of ITS in terms of FIT, and our recommendations for the application of the FIT to different levels of IMI. It can be seen that regardless of IMI level, we recommend routine use of error sensitive feedback (FIT corrective feedback level III); the evidence regarding its benefits are fairly clear, and it is relatively straightforward to implement. With regard to the other factors, the evidence is less clear, and so we present a range of potentially beneficial levels. With regard to macro-sequencing, it should be remembered that Levels I – IV are not mutually exclusive, and can be combined in various ways. Table 9 displays all the FIT levels in one place. Along with Table 8, FIT can be used as guidance when designing or procuring instructional technology.

Table 8. ***Typical FIT Levels Associated With ITS, and Recommended for Different Levels of IMI***

FIT Levels	ITS	Level 1 IMI	Level 2 IMI	Level 3 IMI	Level 4 IMI
Corrective Feedback	III - IV	NA*	III	III	III
Support	I - IV	I	0 - III	0 - III	0 - III
Micro-sequencing	III	NA*	I - II	0 - III	0 - IV
Macro-sequencing	NA*	0 - III	0 - IV	0 - IV	0 - IV

* Not Applicable

Table 9. ***FIT Summary Table***

CORRECTIVE FEEDBACK

Level 0	No explicit item-level feedback – only summary score
Level I	Minimal feedback (item accuracy information)
Level II	Correct answer or explanation of correct answer
Level III	Error-sensitive feedback
Level IV	Context-aware feedback

SUPPORT

Level 0	No support
Level I	Fixed hints on request (problem determined); other fixed sources of information (e.g., glossary) where student initiates access
Level II	Locally-adaptive hints, prompts, or pumps a. on request b. triggered
Level III	Context-aware adaptive hints, prompts, or pumps (True Scaffolding) a. on request b. triggered
Level IV	Same as Level III, with interactive dialog

MICRO-SEQUENCING

Level 0	Recycling
Level I	Supplemental Remediation
Level II	Supplemental Remediation Levels
Level III	Adaptive Content
Level IV	Real-time Adaptation

MACRO-SEQUENCING

Level 0	Fixed sequence (one version only)
Level I	Student choice or hybrid choice/fixed
Level II	Test-out
Level III	Adapted Ahead (multiple versions)
Level IV	Adapted During (multiple versions)

Another use of FIT can be to guide future research. We have already mentioned that evidence comparing the effectiveness of different FIT levels is lacking. This is because research has tended to examine the impact of adaptive interventions in a less than systematic way (varying multiple factors at a time). For example, while ITS have been shown to produce better learning outcomes than other less sophisticated methods of instructional technology, they tend to combine the use of error-sensitive feedback and adaptive content micro-sequencing. The effectiveness of ITS has been compared to methods that do not employ these techniques. Consequently, it is not possible to tell whether the relatively superior effectiveness of ITS is due to the feedback method or the micro-sequencing method (or both). We suggest that systematic research holding the method of feedback constant and varying the level of micro-sequencing is called for. It may be that error-sensitive feedback combined with supplemental remediation or supplemental remediation levels can be just as effective, and easier to implement, than adaptive content. As another example, Mitrovic et al. (2013) demonstrated that adaptive positive and negative feedback messages produced faster learning compared to adaptive negative feedback only; their finding demonstrates that positive feedback can be beneficial, but does not require the interpretation that the positive feedback needs to be adaptive. This is because they did not include a condition with non-adaptive positive feedback.

In conclusion, FIT is a framework that describes the types of instructional decisions that an adaptive learning environment can make. For each instruction decision, FIT lays out a continuum of adaptation that roughly corresponds to the level of sophistication required to support those decisions. Some levels require adaptive tactics that are easier and less resource-intensive to implement, others require more complex student data to allow for contextually adaptive hints, prompts, and feedback and performance adaptive remediation and sequencing. These adaptive tactics are more complex and resource-intensive to implement. One aim of FIT is to help developers of training technology understand the range of possibilities in designing adaptive learning environments. Another aim of FIT is to guide future research. We acknowledge that FIT does not address all the types of decisions a system could make; however, it does describe different possibilities for implementing adaptive decisions and can be used to help designers select the most effective options in keeping with their resources. Therefore, this framework should help Army Leadership and training developers make better decisions when designing adaptive learning environments.

REFERENCES

Aleven, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science, 26*, 147-179.

Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R. M. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research, 73*, 277-320.

Azevedo, R., & Bernard, R. M. (1995). A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research, 13*, 111-127.

Azevedo, R., & Hadwin, A. F. (2005). Introduction to special issue: Scaffolding self-regulated learning and metacognition: Implications for the design of computer-based scaffolds. *Instructional Science, 33,* 367-379.

Behrens, J. T., Mislevy, R. J., DiCerbo, K. E., & Levy, R. (2010). *An evidence centered design for learning and assessment in the digital world.* (CRESST Report 778). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Billings, D. R. (2012). Efficacy of adaptive feedback strategies in simulation-based training. *Military Psychology, 24*, 114-133.

Bloom, B. S. (1976). *Human characteristics and school learning*. New York: McGraw-Hill.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher 13*, 4–16.

Brusilovsky, P., & Vassileva, J. (2003). Course sequencing techniques for large-scale web-based education. *Conference on Intelligent Tutoring Systems.* Springer-Verlag: Berlin.

Clarebout, G., & Elen, J. (2009). Benefits of inserting support devices in electronic learning environments. *Computers in Human Behavior, 25*, 804-810.

Corbalan, G., Kester, L., & van Merriënboer, J. J. G. (2008). Selecting learning tasks: Effects of adaptation and shared control on learning efficiency and task involvement. *Contemporary Educational Psychology, 33*, 733-756.

Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4,* 253-278.

Corcoran, T., Mosher, F. A., & Rogat, A. (2009). Learning progressions in Science: An Evidence-based approach to reform. *Consortium for Policy Research in Education* (Report # RR-63). Retrieved from http://www.cpre.org/images/stories/cpre_pdfs/lp_science_rr63.pdf.

desJardins, M., Ciavolino, A., Deloatch, R., & Feasley, E. (2011). Playing to program: An intelligent programming tutor for RUR-PLE. *Proceedings of the Second AAAI Symposium on Educational Advances in Artificial Intelligence*, 9–10 August, San Francisco, CA.

Desmarais, M. C., & Baker, R. S. J. d. (2012). A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction*, *22*, 9–38.

Domeshek, A., Durlach, P. J., & Bratt, E. O. (2010). A framework for defining and experimenting with adaptation in on-line training. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*, Orlando, FL.

Durlach, P. J., & Ray, J. M (2011). *Designing adaptive instructional environments: Insights from empirical evidence*. (Technical Report 1297). Arlington, VA: U. S. Army Research Institute for the Behavioral Social Sciences.

Durlach, P. J., & Spain, R. D. (2012). Framework for instructional technology. In V. G. Duffy (Ed.), *Advances in Applied Human Modeling and Simulation* (pp. 222–231). Boca Raton, FL: CRC Press.

Durlach, P. J., Wansbury, T. G., & Wilkinson, J. G. (2008). Cultural awareness and negotiation skills training: Evaluation of a prototype semi-immersive system. *Army Science Conference*, December, Orlando, FL.

Dutke, S., & Reimer, T. (2000). Evaluation of two types of online help for application software. *Journal of Computer Assisted Learning*, *16*, 307-315.

Fautua, D. T., Schatz, S., Reitz, E., & Killilea, J. (2012). Joint continuum of eLearning: Implementing engaging, effective, and meaningful military eLearning. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*, 3–6 Dec., Orlando, FL.

Forbes-Riley, K., & Litman, D. (2011). Designing and evaluating a wizarded uncertainty-adaptive spoken dialogue tutoring system. *Computer Speech and Language*, *25*, 105-126.

Fossati, D. (2009). *Automatic modeling of procedural knowledge and feedback generation in a computer science tutoring system* (Doctoral Dissertation). University of Illinois, Chicago. Retrieved from http://www.fossati.us/papers/ilist-phdthesis.pdf

Graesser, A. C., Jeon, M. & Dufty, D. (2008). Agent technologies designed to facilitate interactive knowledge construction. *Discourse Processes*, *45*, 298-322.

Guo , Y., Heffernan, N. T., & Beck, J. E. (2008). Trying to reduce bottom-out hinting: Will telling student how many hints they have left help? *Proceeding of the 9th International Conference on Intelligent Tutoring Systems* (pp. 774–778). Berlin: Springer-Verlag

Guskey, T. R., & Pigott, T. D. (1988). Research on group-based mastery learning programs: A meta-analysis. *Journal of Educational Research*, *81*, 197–216.

Hattie J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*, 81–112.

Hays, R. T. (2005). *The effectiveness of instructional games: A literature review and discussion*. (Technical Report 2005-004). Orlando, FL: Naval Air Warfare Center Training Systems Division.

Hays, M. J., Kornell, N., & Bjork, R. A. (2013). When and why a failed test potentiates the effectiveness of subsequent study. *Journal of Experimental Psychology: Learning, Memory & Cognition, 39*, 290-296.

Heffernan, N. T., Koedinger, K. R., & Razzaq, L. (2008). Expanding the model-tracing architecture: A 3rd generation intelligent tutor for algebra symbolization. *International Journal of Artificial Intelligence in Education, 18*, 153-178.

Jaehnig, W., & Miller, M. L. (2007). Feedback types in programmed instruction: A systematic review. *The Psychological Record*, *57*, 219-232.

Johnson, W. L., Friedland, L., Schrider, P., Valente, A., & Sheridan, S. (2011). The Virtual cultural awareness trainer (VCAT): Joint Knowledge Online's (JKO's) solution to the individual operational culture and language training gap. Paper presented at ITEC, 10–12 May, London.

Kali, Y., & Linn, M. C. (2008). Technology-enabled support strategies for inquiry learning. In J. M. Spector, M. D. Merrill, J. Van Merrienboer, & M. P. Driscoll (Eds.). *Handbook of Research on Educational Communications and Technology (3rd Edition*, pp. 145-161). New York: Lawrence Erlbaum Associates.

Kalyuga, S., & Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction. *Journal of Educational Psychology*, *96*, 558-568.

Kim, J., Hill, R., Durlach, P., Lane, H., Forbell, E., Core, M., Marsella, S., Pynadath, D., & Hart, J. (2009). BiLAT: A game-based environment for practicing negotiation in a cultural context. *Special Issue on Ill-Defined Domains: International Journal of AI in Education*, *19*, 289-308.

Kluger, A., & DeNisi (1998). Feedback interventions: Toward the understanding of a double-edged sword. *Current Directions in Psychological Science, 7*, 67-72.

Koedinger, K., McLaughlin, E., & Stamper, J. (2012). Automated student model improvement. *Proceedings of the 5th International Conference on Educational Data Mining (EDM 2012)*. Chania, Greece, pp. 17-24.

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review, 1*, 279-308.

Kulik, C. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research*, *60*, 265–299.

Landsberg, C. R., Astwood, R. S., Van Buskirk, W, L., Townsend, L, N., Steinhauser, N. B., & Mercado, A. D. (2012). A review of adaptive training system techniques. *Military Psychology*, *24*, 2012.

Levchuk, G., Shebilske, W., & Freeman, J. (2012). A model-driven instructional strategy: The benchmarked experiential system for training (BEST). In P. J. Durlach & A. M. Lesgold (Eds.) *Adaptive technologies for training and education*. (pp. 303–317). Cambridge, UK: Cambridge University Press.

Luckin, R., & du Boulay, B. (1999). Ecolab: The development and evaluation of a Vygotskian design framework. *International Journal of Artificial Intelligence in Education*, *10*, 198-220.

Mitrovic, A., Ohlsson, S., & Barrow, D. K. (2013). The effect of positive feedback in a constraint-based intelligent tutoring system. *Computers & Education, 60*, 264-272.

Munro, A., Fehling, M. R., & Towne, D. M. (1985). Instruction intrusiveness in dynamic simulation training. *Journal of Computer-Based Instruction, 12*, 50-53.

Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. J. G. van Merriënboer, & M. Driscoll (Eds.), *Handbook of research on Educational communications and technology, (3rd edition,* pp. 125-144). New York: Lawrence Erlbaum Associates.

Park, O., & Lee, J. (2004). Adaptive instructional systems. In D. H. Jonassen (Ed.), *Handbook of Research on Educational Communications and Technology* (pp. 651–684). Mahwah, N. J.: Lawrence Erlbaum Associates.

Pashler, H., McDaniel, M., Rohrer, D. & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest*, *9*, 105–119.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences*. *13*, 423-451.

Perrin, B. M., Dargue, B. W., & Banks, F. Z. (2003). Dynamically adapting content delivery: An effectiveness study and lessons learned. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC).* Orlando, FL.

Renkl, A., Hilbert, T., & Schworm, S. (2009). Example-based learning in heuristic domains: A cognitive load theory account. *Educational Psychology Review, 21*, 67-78.

Roll, I., Aleven, V., McLaren, B. M., & Koedinger, K. R. (2011). Improving students' help-seeking skills using meta-cognitive feedback in an intelligent tutoring system. *Learning and Instruction, 21*, 267-280.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119-144.

Salas, E., Burke, S. C., & Cannon-Bowers, J. A. (2000). Teamwork: Emerging principles. *International Journal of Management Reviews, 2*, 339-356.

Salden, R., Aleven, V., Renkl, A., & Schwonke, R. (2009). Worked examples and tutored problem solving: Redundant or synergistic forms of support? *Topics in Cognitive Science*, *1*, 203-213.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*, 207-217.

Schwonke, R., Renkl, A., Salden, R., & Aleven, V. (2011). Effects of different ratios of worked solution steps and problem solving opportunities on cognitive load and learning outcomes. *Computers in Human Behavior, 27*, 58-62.

Shute, V. J. (1993). A comparison of learning environments: All that glitters... In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 47-74). Hillsdale, NJ: Lawrence Erlbaum Associates.

Shute, V. J., & Kim, Y. J. (2013). Formative and stealth assessment. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology (4th Edition).* New York: Taylor & Francis Group. Retrieved on 26 Nov, 2012 from http://myweb.fsu.edu/vshute/pdf/formative%20stealth%20assessment%20FINAL.pdf.

Shute, V. J., & Psotka, J. (1996). Intelligent tutoring systems: Past, present, and future. In D. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 570-600). New York: Macmillan. Retrieved from http://myweb.fsu.edu/vshute/pdf/shute 1996_d.pdf

Shute, V. J., & Zapata-Rivera, D. (2008). Adaptive technologies. In J. M. Spector, D. Merrill, J. van Merriënboer, &M. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd edition, pp. 277–294). New York: Lawrence Erlbaum Associates.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257–285.

TRADOC (2010). *Distributed learning education and training products delivery order for new/redesign of dL IMI and blended distributed learning.* Retrieved from http://www.atsc.army.mil/itsd/imi/BLDO/DLETPDOTemplates.asp.

TRADOC (2011). *The U. S. Army Learning Concept for 2015.* (TRADOC Pamphlet 525-8-2). Retrieved from http://www.tradoc.army.mil/tpubs/pams/tp525-8-2.pdf.

Triantafillou, E., Pomportsis, A., Demetriadis, S., & Georgiadou, E. (2004). The value of adaptivity based on cognitive style: An empirical study. *British Journal of Educational Technology, 35,* 95-106.

Tseng, J. C. R., Chu, H., Hwang, G., & Tsai, C. (2008). Development of an adaptive learning system with two sources of personalization information. *Computers & Education, 51,* 776-786.

Tsiriga, V., & Virvou, M. (2004). Evaluating the intelligent features of a web-based intelligent computer assisted language learning system. *International Journal on Artificial Intelligence Tools, 13*, 411-425.

Vandewaetere, M., Desmet, P., & Clarebout, G. (2011). The contribution of learner characteristics in the development of computer-based adaptive learning environments. *Computers in Human Behavior, 27,* 118-130.

Van Dijk, D., & Kluger, A. (2011). Task type as a moderator of positive/negative feedback effects on motivation and performance: A regulatory focus perspective. *Journal of Organizational Behavior. 32,* 1084-1105.

VanLehn, K. (1982). Bugs are not enough: An analysis of systematic subtraction errors. *Journal of Mathematical Behavior, 3,* 3-71.

VanLehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education, 16,* 227-265.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems and other tutoring systems. *Educational Psychologist, 46,* 4, 197-221.

Vygotsky, L.S. (1978). *Mind and society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wood, H., & Wood, D. (1999). Help seeking, learning and contingent tutoring. *Computers and Education*, *33,* 153–169.

Woolf, B. P. (2009). *Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning*. Burlington MA: Morgan Kaufmann Publishers.

ACRONYMS

AAR	After Action Review
ALC 2015	Army Learning Concept 2015
ATI	Aptitude Treatment Interaction
DOD	Department of Defense
ELO	Enabling Learning Objective
FIT	Framework for Instructional Technology
IMI	Interactive Multimedia Instruction
ITS	Intelligent Tutoring System(s)
TLO	Terminal Learning Objective
TRADOC	Training and Doctrine Command
VCAT	Virtual Cultural Awareness Trainer